Coding of Natural Scenes in Primary Visual Cortex

Michael Weliky,* József Fiser, Ruskin H. Hunt,
David N. Wagner
Department of Brain and Cognitive Sciences
Meliora Hall
University of Rochester
Rochester, New York 14627

Summary

Natural scene coding in ferret visual cortex was investigated using a new technique for multi-site recording of neuronal activity from the cortical surface. Surface recordings accurately reflected radially aligned layer 2/3 activity. At individual sites, evoked activity to natural scenes was weakly correlated with the local image contrast structure falling within the cells' classical receptive field. However, a population code, derived from activity integrated across cortical sites having retinotopically overlapping receptive fields, correlated strongly with the local image contrast structure. Cell responses demonstrated high lifetime sparseness, population sparseness, and high dispersal values, implying efficient neural coding in terms of information processing. These results indicate that while cells at an individual cortical site do not provide a reliable estimate of the local contrast structure in natural scenes, cell activity integrated across distributed cortical sites is closely related to this structure in the form of a sparse and dispersed code.

Introduction

The response of striate cortical cells to simple stimuli, such as isolated bars or gratings presented within their classical receptive field, can be described as performing a spatially restricted orientation and spatial frequency filtering of the visual field (DeValois et al., 1982; Hubel and Wiesel, 1959). This response behavior shares many properties with optimal linear filters, such as originally proposed by Gabor (1946), that encode any complex visual scene by the magnitude of local Fourier, or 2D spatial frequency, components. However, such filter properties cannot fully account for striate cortical cell responses to more complex stimuli. For example, a cell's response to a stimulus of optimal orientation and spatial frequency can be suppressed or facilitated by additional stimuli placed either within (DeAngelis et al., 1992) or outside the classical receptive field (Blakemore and Tobin, 1972; Gilbert and Wiesel, 1990). This response modulation can vary from one cell to another according to the relative orientation (Nelson and Frost, 1978), spatial frequency (Chao-Yi and Wu, 1994; DeValois and Tootell, 1983), location (Walker et al., 1999), contrast (Levitt and Lund, 1997; Polat et al., 1998), or brightness (Rossi et al., 1996) of the compound stimuli and cannot be

*Correspondence: weliky@cvs.rochester.edu

predicted from the response to simple stimuli presented within the cell's receptive field.

Since it has proven difficult to characterize the coding of complex stimuli at the level of individual cells, here we test a simple population-coding model in which cell activity is integrated across distributed cortical sites having retinotopically overlapping receptive fields. We ask whether this integrated activity carries the information of the local contrast structure of natural scenes as defined by the local Fourier components falling in the cells' classical receptive fields. Such a finding would not only demonstrate that this information is represented as a distributed population code, but would also help to establish a relationship between cortical coding and the large body of data collected on cell responses in the last decades with simple stimuli such as bars and gratings.

In our study, cell activity was simultaneously recorded from multiple cortical sites of the striate cortex of the anesthetized ferret using a new multielectrode surface recording technique. This method allowed high density mapping of cell activity at many recording sites distributed across a large area of cortex. At each surface recording site, we obtained multiunit responses with properties similar to a radially aligned layer 2/3 recording, without risk of damaging the brain by large numbers of penetrating electrodes. The classical receptive field tuning properties of cells at each site, as measured by simple sine wave gratings, were used to predict responses to natural scenes at three different levels of encoding: (1) individual cortical sites, (2) simultaneously recorded multiple cortical sites, and (3) the weighted retinotopic average of population activity. Our results demonstrate a weak correlation between cell activity and the local contrast structure of natural scenes when computed at individual cortical sites across many different images. This correlation is stronger but still fairly weak when computed across multiple cortical sites for any individual image. However, a simple population code, derived from activity integrated across cortical sites having retinotopically overlapping receptive fields, is highly correlated with the local contrast structure. This high correlation between measured population responses and the underlying contrast structure does not change significantly when only the classical receptive field is stimulated or the stimulus is extended to a large field around the classical receptive field.

The efficiency of coding in the striate cortex was also characterized. The measured responses showed high sparseness compared to not only the predicted responses based on classical Gabor filter-based coding, but also compared to other coding schemes proposed earlier on a theoretical basis. Thus, the striate cortex generates a sparse and dispersed code that is highly efficient for storing information and associative learning (Treves and Rolls, 1991). This code represents the local contrast structure of natural scenes in a distributed manner that could be extracted by higher cortical areas.
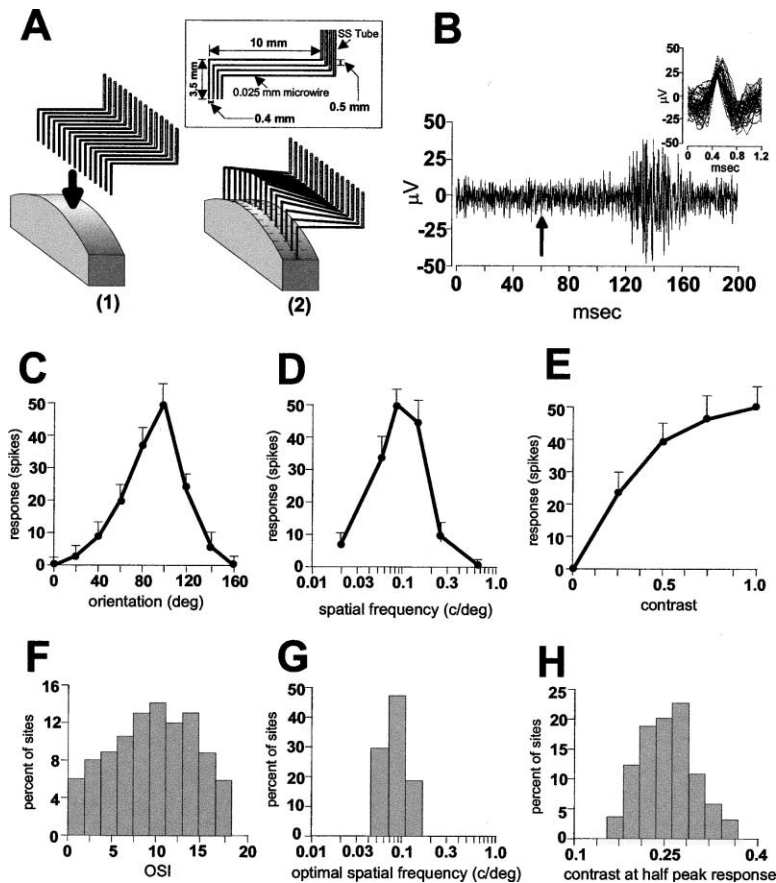
Figure 1. Multi-Site Cortical Surface Recording

(A) A single row of 15 flexible electrodes demonstrates how the array conforms to the curved shape of the cortical surface. Electrode spacing is 0.6 mm. A cross-section through the array is shown in the inset. Electrode rows are separated by 0.4 mm. Successive rows are shifted downward by 0.5 mm to prevent electrodes in one row from hitting those in another row as the electrodes move up and down. Microwire electrodes are held within small diameter stainless steel (SS) tubing.

(B) Evoked neural activity at a single surface recording site in response to the presentation of a flashed natural scene image. The image was flashed on the display screen at the time shown by the arrow. A burst of small amplitude action potentials is seen about 60 ms after the stimulus onset. The inset shows an expanded view of waveforms of these action potentials extracted from the trace.

(C–E) Receptive field response properties assessed by cortical surface recording at one site. (C) Orientation tuning. The OSI of this site is 9.5. (D) Spatial frequency tuning. (E) Contrast response.

(F–H) Distribution of receptive field tuning properties assessed by cortical surface recording at all sites across all animals. (F) OSI values. (G) Spatial frequency tuning. (H) Contrast evoking half-peak response.

Results

Multielectrode Cortical Surface Recording

Neuronal activity was simultaneously recorded from up to 60 separate sites across the surface of an 8.4 by 1.2 mm area of visual cortex using an array of flexible 25 μm diameter microwire electrodes (Figure 1A). The impedance of the electrodes ranged from 110 to 130 KΩ. The flexible electrode design permitted the array to conform to the curvature of the cortex, allowing electrodes to move up and down with normal pulsation while maintaining stable contact with the cortical surface. Although we can not rule out the possibility that the electrodes may have penetrated slightly into the cortex, electrodes were placed on the cortical surface with light force that caused, at most, only slight local dimpling. In control experiments where an electrode was first placed on the surface and then pushed into the cortex, there was significant dimpling as the electrode was advanced before it pierced the pia. In addition, there was an immediate and noticeable increase in the amplitude of spontaneous and driven activity when the electrode entered the cortex.

Evoked by flashed sine wave gratings, small boxes, or natural scene images, low amplitude signals were recorded from the cortical surface (Figure 1B). The amplitude of these signals typically ranged between 15μV and 100μV, and the waveform showed the typical shape of action potentials (see Figure 1B, inset). While it was necessary to keep the cortical surface moist, surface recordings of sufficient quality to observe these signals could not be obtained if there was a large amount of fluid buildup. Since the signal was bandpass filtered over the same range (600–6000 Hz) used by previous studies to segregate action potentials from local field potentials (DeAngelis et al., 1998; Fries et al., 1997; Logothetis et al., 2001), we anticipate that this recorded surface activity reflects primarily action potentials and not local field potentials, which are typically low-pass filtered below 100 Hz. However, we cannot completely rule out the possibility that subthreshold voltage changes may also contribute to the recorded signal.

Receptive Field Response Properties Assessed by Cortical Surface Recording

Based on our surface recordings, we carried out a complete characterization of neural responses in the primary cortex by determining the receptive field position and size, orientation and spatial frequency tuning, and contrast response at individual recording sites (Figures 1C–1H). Recordings were obtained from an average of 45.6 ± 5.1 electrodes across all experiments (n = 5 animals). The size and location of receptive fields at all electrodes were simultaneously mapped using a standard reverse correlation technique with small white squares. Receptive fields ranged in size from 8° to 20° in diameter located within the central 30° of the visual field. The aggregate visual field covered by receptive fields at all electrodes ranged from 35° to 50° in elevation by 20° to 35° in azimuth.

Orientation tuning, spatial frequency tuning, and contrast response functions were determined at each recording site using flashed sine wave gratings. Responses to sine wave gratings presented in 90 degrees of phase shift were not statistically different (p > 0.1, t test), suggesting that our recordings were either dominated by responses from complex cells or that simple cell responses pooled to produce the phase independence. To quantify the cell responses, an orientation selectivity index (OSI) was computed for each site according to the measure defined in Weliky and Katz (1997). The mean OSI was 8.94 ± 0.15 (Figures 1C and 1F). The cell responses had a mean peak spatial frequency tuning of 0.095 ± 0.03 cycle/deg and a mean bandwidth (half height) of 3.2 ± 0.53 octaves (Figures 1D and 1G). Contrast response functions at all sites exhibited nonlinear saturation at high contrasts, as previously described for primary visual cortical cells (Albrecht and Hamilton, 1982) (Figures 1E and 1H).

These characteristics of cells are in very good agreement with earlier results reported with studies using penetrating electrodes in ferrets. The receptive fields in the ferret are large compared to cats or monkeys. Earlier studies in the ferret found receptive fields varying in size from 10° to 20° (White et al., 1999), which is close to what we found with our method. The reported mean OSI value with penetrating electrodes was previously found to be 9.6, which is close to the mean value reported here (Weliky and Katz, 1997).

Surface Recordings Reflect Layer 2/3 Neural Responses

Simultaneous multi-unit recordings from radially aligned surface and penetrating electrodes were obtained with a three-unit recording array, each unit consisting of a double electrode (Figure 2A), with 600 μm spacing between units (n = 12 sites). In each unit, the tips of the two electrodes were shifted by 250 μm in depth with respect to each other, allowing one of the electrodes to penetrate the cortex while the other one only touched the surface nearby the point of penetration. The recordings from the penetrating electrode were analyzed by two standard techniques: with voltage threshold discrimination for layer 2/3 multi-unit recording data, and with time-amplitude window discrimination for layer 2/3 single unit recordings. The recordings from the surface electrode were analyzed with voltage threshold discrimination. Because well-isolated single units could not be extracted from all penetrating recording sites, we have analyzed data only from double electrode recordings that yielded well-isolated single units at the penetrating site (n = 8 sites). Cell activity recorded at the cortical surface shared highly similar response properties with responses obtained by standard penetrating recordings at a radially aligned site within layer 2/3. Waveform amplitudes were on the average 2.1 times higher with penetrating recordings than with surface recordings, while the mean RMS noise for surface (4.69 ± 0.49) and penetrating (4.81 ± 0.38) recordings were statistically similar (p > 0.16, t test). The receptive field positions and sizes based on reverse correlation were virtually identical (r = 0.96 ± 0.05 between surface and multi-unit, r = 0.93 ± 0.08 between surface and single unit) (Figures 2D–2F).

The 2D receptive field mapping technique, utilizing individually flashed square stimuli, was further used to assess whether surface electrodes collected signal over a wider cortical area than penetrating electrodes. First, we determined the retinotopic location of the flashed square that evoked the strongest response at each surface and radially aligned penetrating electrode. We then computed the mean response to flashed squares located at varying retinotopic distances from the site evoking the peak response (Figure 2B). If surface electrodes collected signal over a wider cortical area than penetrating electrodes, the responses at surface recording sites should drop with a much shallower slope than the responses at penetrating recording sites as a function of increasing retinotopic distance. The mean normalized response to flashed squares located at increasing retinotopic distances from the site of peak response was statistically indistinguishable between surface and penetrating multi-unit/single-unit recordings (p > 0.1, t test). In a second test, we computed the correlation between receptive fields recorded at pairs of surface recording sites separated by 0.6 mm or 1.2 mm. The correlation between receptive fields was reduced with increasing separation between sites. These correlations were compared to those computed independently for pairs of penetrating multi-unit and pairs of single-unit recordings (Figure 2C). Correlations were not statistically different for surface recordings and penetrating multi-unit/single-unit recordings at corresponding site separations (p > 0.1, t test). If surface recordings reflect a signal from a wider cortical area, it would have been expected that correlations between surface recording sites would drop with a shallower slope. Together, these tests demonstrate that the surface and penetrating recordings collected signals over equivalent cortical areas.

There was a similarly high correlation between orientation, spatial frequency, and contrast tuning curves recorded with surface and penetrating electrodes at radially aligned sites (Figure 3). For orientation tuning curves, the correlation between the OSI of surface and penetrating multi-unit recordings was r = 0.93 ± 0.09 and r = 0.85 ± 0.11 between surface and penetrating single-unit recordings. The correlation between peak orientation tuning of surface and penetrating multi-unit recordings was r = 0.96 ± 0.03 and r = 0.96 ± 0.03 between surface and penetrating single-unit recordings. For spatial frequency tuning curves, the correlation between the bandwidth (half height) of surface and penetrating multi-unit recordings was r = 0.92 ± 0.04 and r = 0.89 ± 0.05 between surface and penetrating single-unit recordings. The correlation between peak spatial frequency tuning of surface and penetrating multi-unit recordings was r = 0.98 ± 0.04 and r = 0.98 ± 0.04 between surface and penetrating single-unit recordings. Contrast tuning showed a similarly high agreement between surface and penetrating recordings, with r = 0.94 ± 0.04 between half height contrast response of surface and penetrating multi-unit and r = 0.91 ± 0.09 between half height contrast response of surface and penetrating single-unit recordings.

Simultaneous cell responses to 116 natural scenes were also collected with radially aligned surface and penetrating electrodes. Figure 4A shows three examples
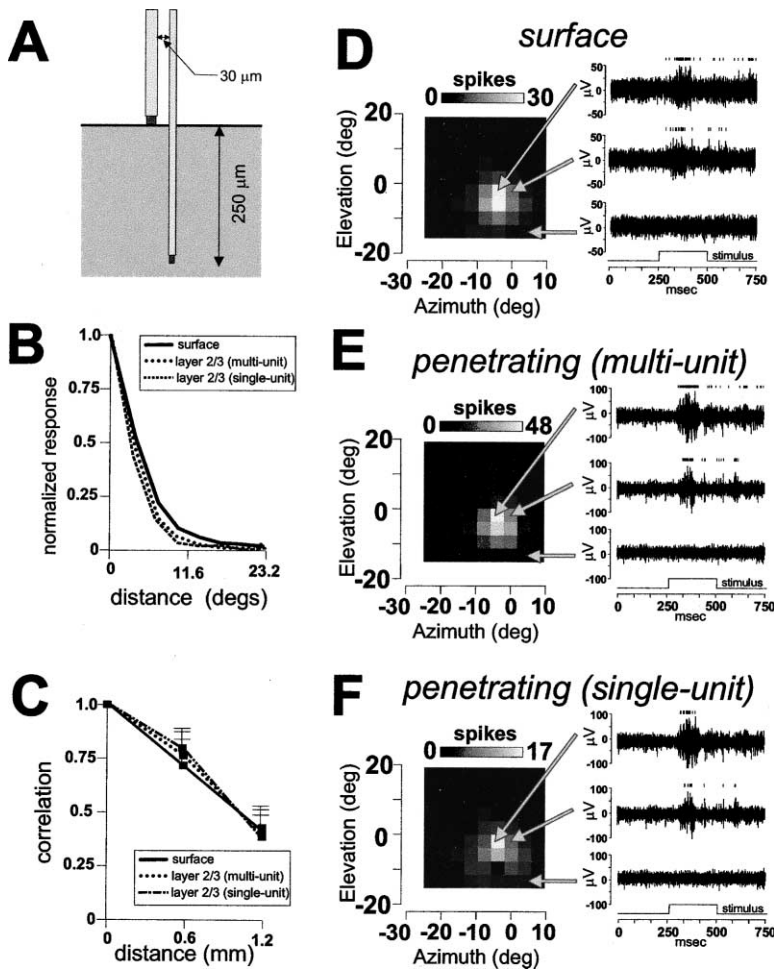
Figure 2. Comparison of Receptive Fields Simultaneously Recorded from Surface and Penetrating Electrodes

(A) Diagram of the double electrode design allowing simultaneous surface and penetrating recordings. Surface and penetrating electrodes are made from 25 μm and 12.5 μm diameter tungsten microwires, respectively.

(B) Mean falloff of evoked responses at individual surface, penetrating multi-unit, and single-unit recording sites to flashed squares.

(C) Mean correlation between receptive fields recorded at pairs of surface recording sites separated by 0.6 or 1.2 mm. These correlations are plotted together with those computed for penetrating multi-unit and single-unit recordings.

(D–F) Plots of receptive fields simultaneously recorded from the surface and a second site 0.25 mm directly below in layer 2/3. Two-dimensional maps show evoked spike activity plotted as a function of the spatial location of individually flashed squares in a 10 by 10 matrix. Grayscale encodes the spike activity response to the presentation of each square. Traces on the right show spike discharges at three receptive field locations indicated by the arrows. Flash onset occurred following a 250 ms baseline period and had a duration of 250 ms. Vertical line segments above each trace show discriminated spikes. (D) Multi-unit surface recording. (E) Multi-unit penetrating layer 2/3 recording. (F) Single-unit penetrating layer 2/3 recording obtained using time/amplitude window discrimination to extract a single spike waveform from the multi-unit traces shown in (E).

of natural scene images with increasing articulation of a sharp contrast edge within the receptive field of one unit, and the corresponding surface and penetrating recordings. While penetrating recordings typically showed higher firing rates than surface recordings, both types of recordings exhibited a similar orderly increase in firing rate to these three images. To quantify this relation, we computed the correlation between surface versus penetrating multi-unit and surface versus penetrating single-unit activity for all 116 images (Figure 4B). In response to natural scenes, we found a similarly high correlation between surface and penetrating recording as with gratings ($r = 0.94 \pm 0.09$, and $r = 0.84 \pm 0.6$ for surface versus multi-unit and surface versus single-unit, respectively).

Using natural scenes, we assessed whether the falloff in response correlation between pairs of surface recording sites was different from that obtained between pairs of penetrating multi-unit or single-unit recordings (Figure 4C). The average correlation between surface recording sites separated by 0.6 and 1.2 mm was $r = 0.65 \pm 0.05$ and $r = 0.45 \pm 0.09$, respectively. The finding that these values are very close to that obtained by individual flashed square stimuli (see Figure 2C) is surprising, since it might be expected that because natural scenes have a different statistical structure (i.e., extended contours and edges), the falloff in correlations

would be more shallow. The correlations for penetrating multi-unit and penetrating single-unit recordings were not statistically different from those computed for surface recordings ($p > 0.1$, t test).

A Cell Response Model Based on Measuring the Contrast Structure within the Classical Receptive Field

To investigate whether responses of cells in the striate cortex to natural scenes encode the 2-dimensional contrast structure of stimuli falling within the classical receptive field, we generated an objective measurement of this information. Any image falling on the retina is a 2D array of intensity values where each pixel is defined by two spatial coordinates and an intensity value (Figure 5A). An equivalent representation of the same image is its Fourier representation, which defines the image in terms of its 2D spatial frequency components: each point defines the amplitude of a 2D sine wave of a specific orientation (angle) and spatial frequency (distance of the point from center) (Figure 5D). The overwhelming majority of cells in the striate cortex respond to sine wave stimuli falling within different restricted ranges of orientations and spatial frequencies (Figure 5B). This means that each cell is sensitive to only a restricted area of the Fourier domain, with a bell-shape weighting around the coordinates defined by the peak orientation
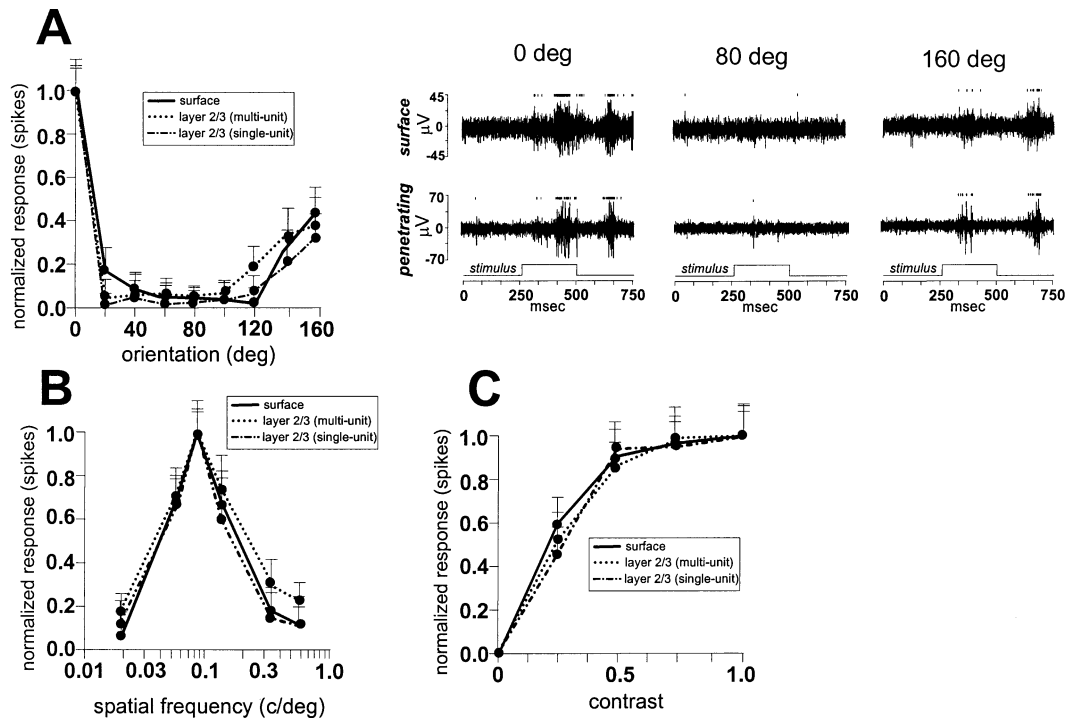
Figure 3. Comparison of Receptive Field Tuning Properties Simultaneously Recorded from Radially Aligned Surface and Penetrating Electrodes

These recordings were obtained using the same double electrode shown in Figure 2A.
(A) Orientation tuning curves obtained simultaneously from multi-unit surface activity, and multi-unit and single-unit layer 2/3 activity. Traces on the right show spike discharges at three different grating orientations of 0°, 80°, and 160°. 90° represents a vertically oriented grating. Grating flash onset occurred following a 250 ms baseline period and had a duration of 250 ms. Vertical line segments above each trace show discriminated multi-unit spikes.
(B) Spatial frequency tuning.
(C) Contrast response function.

and spatial frequency tuning of the cell (Figure 5E). Multiplying the bell-shape Fourier representation of the cell's tuning function with the Fourier representation of the image (Figure 5F) and then inverting back the result into the spatial domain by the inverse Fourier transform gives, by definition, all the contrast components in the image that fall in the cell's preferred range of orientations and spatial frequencies (Figure 5C). The filtered image has to be masked with the receptive field of the cell to obtain the 2D contrast components of the image that fall within the cell's classical receptive field (Figures 5G–5I).

Once we obtained the contrast components within the cells' classical receptive field, as well as orientation and spatial frequency tuning, we needed to select a measurement that gives a good prediction of the cell's response when it is stimulated with the given contrast component. Since in our recordings the change of phase did not alter the responses, the energy model of complex cell responses was an appropriate model. The simplest approximate measure of energy in an image is the contrast of the image. We tested two contrasts, the Michelson contrast and the RMS contrast. Since the contrast structures relevant to the ferret cortical cells were restricted to a narrow low-frequency band, the two measures gave indistinguishable results from the point of view of our conclusions. For computational simplicity, we present our results using the Michelson contrast.

Finally, the obtained contrast value was scaled by the nonlinear contrast function derived experimentally from responses to sine wave gratings. At each recording site, this final value was our estimate of the response to the original image based on the 2D contrast components that fell within the classical receptive field; hence, we will refer to these values as expected or predicted responses.

In order to demonstrate the validity of the model on simple stimulus inputs, during an experiment we inserted trials containing our full set of sine wave gratings (see Experimental Procedures) masked to the measured receptive field of one or two recording sites (n = 7 sites). We then tested how well the model predicted the responses to these stimuli. There was a high correlation between orientation ($r = 0.86 \pm 0.15$), spatial frequency ($r = 0.89 \pm 0.17$), and contrast tuning curves ($r = 0.87 \pm 0.11$) obtained from the model and these measured responses. Thus, the model reproduced the experimentally measured cell responses to simple sine wave grating stimuli very well.

Response to Natural Scenes at Individual and across Multiple Cortical Sites

Evoked cell activity was recorded simultaneously at all recording sites in response to flashed large-field images of 116 natural scenes (n = 5 animals). We first computed the correlation between measured and expected re-
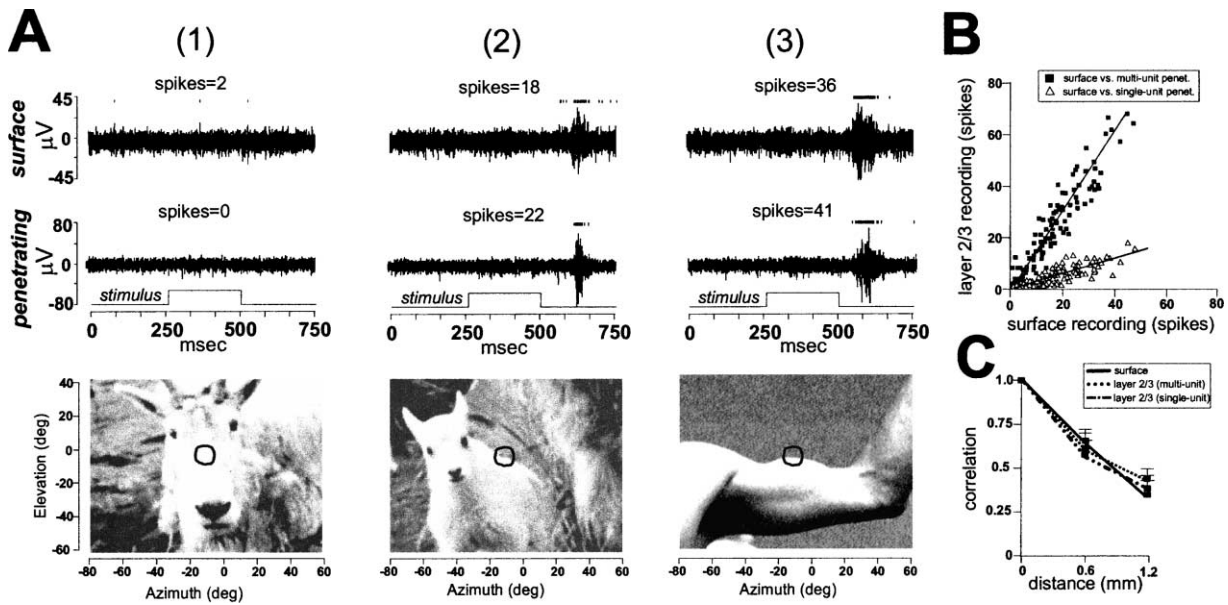
Figure 4. Comparison of Responses to Natural Scenes Simultaneously Recorded from Surface and Penetrating Electrodes

These recordings were obtained using the same double electrode shown in Figure 2A.

(A) Upper traces show increasing spike discharge at both surface and penetrating recording sites in response to three different natural scenes (1–3). Vertical line segments above each trace show discriminated multi-unit spikes. Image flash onset occurred following a 250 ms baseline period and had a duration of 250 ms. The black circle overlaid upon each image shows the location of the classical receptive field.

(B) Scatter plots of simultaneously recorded surface and layer 2/3 multi-unit and single-unit activity in response to 116 flashed natural scenes. Multi-unit surface activity is plotted as a function of multi-unit layer 2/3 activity (black squares, $r = 0.93$) and single-unit layer 2/3 activity (white triangles, $r = 0.84$).

(C) Plot of the correlation in evoked spike discharge rate between simultaneously recorded pairs of (1) surface or (2) penetrating sites separated by 0.6 mm and 1.2 mm in response to 116 natural scenes. Solid line, surface; dotted line, penetrating multi-unit; dashed line, penetrating single unit.

sponses at individual sites to all images. At less than 3% of the recording sites, measured and expected responses were well correlated above $r = 0.6$ (mean $r = 0.17 \pm 0.13$) (Figures 6A and 6B). This indicates that cell responses are not closely related to the magnitude of 2D image components falling within the classical receptive field and orientation/spatial-frequency tuning. Thus, striate cortical cell responses to natural scenes are not consistent with the simple local 2D-oriented spatial frequency filter energy model.

One potential reason for the inconsistency between measured and expected responses is that spike extraction, by voltage thresholding, is an inappropriate method for analyzing the multi-unit signal recorded from the cortical surface. Errors may arise because the spikes from different neurons that fire within a very short time interval temporally sum to produce a single spike waveform crossing the voltage threshold. In order to test this possibility, we full-wave rectified and then integrated the recorded signal. This method will capture any temporally summated signal arising from different neurons. Following quantification of the cortical surface signal using this method, expected responses were computed exactly as for the voltage thresholded signal. Still, a very weak correlation between expected and measured responses was found at each recording site when computed across all images. The distribution of these correlation values was statistically indistinguishable from that obtained by voltage thresholding the signal ($p > 0.05$, t test) (Figure

6B). This verifies that our results are not dependent upon using voltage detection to extract multi-unit activity from the surface recorded signal.

A second potential reason for the weak correlation between measured and expected responses is that we underestimated the size of the receptive field. It has been demonstrated that depolarizing subthreshold inputs can be evoked over an area at least 2.5 times larger than the classical receptive field as defined by the region in which a stimulus elicits spike discharge (Bringuier et al., 1999). These subthreshold inputs could play a role in driving cell activity. Therefore, we recomputed the expected model responses using receptive field masks that were two, four, and six times the diameter of the measured receptive field spike discharge region. With 2-fold larger diameter receptive fields, there was a small increase in the mean correlation between measured and expected responses (mean $r = 0.22 + 0.19$) ($p < 0.05$, t test) (Figure 6B). The mean correlation was again only slightly increased with 4-fold (mean $r = 0.26 + 0.15$) and 6-fold (mean $r = 0.28 + 0.15$) larger diameter receptive fields ($p < 0.05$, t test) (Figure 6B). This test assumes that the structure of the surround is uniform. Typically, when the surround and center stimuli have similar orientation, stimuli within only a localized region of the surround will suppress responses to stimuli within the classical receptive field (Walker et al., 1999). Thus, the recomputation of expected responses using larger receptive field masks may have overestimated the re-
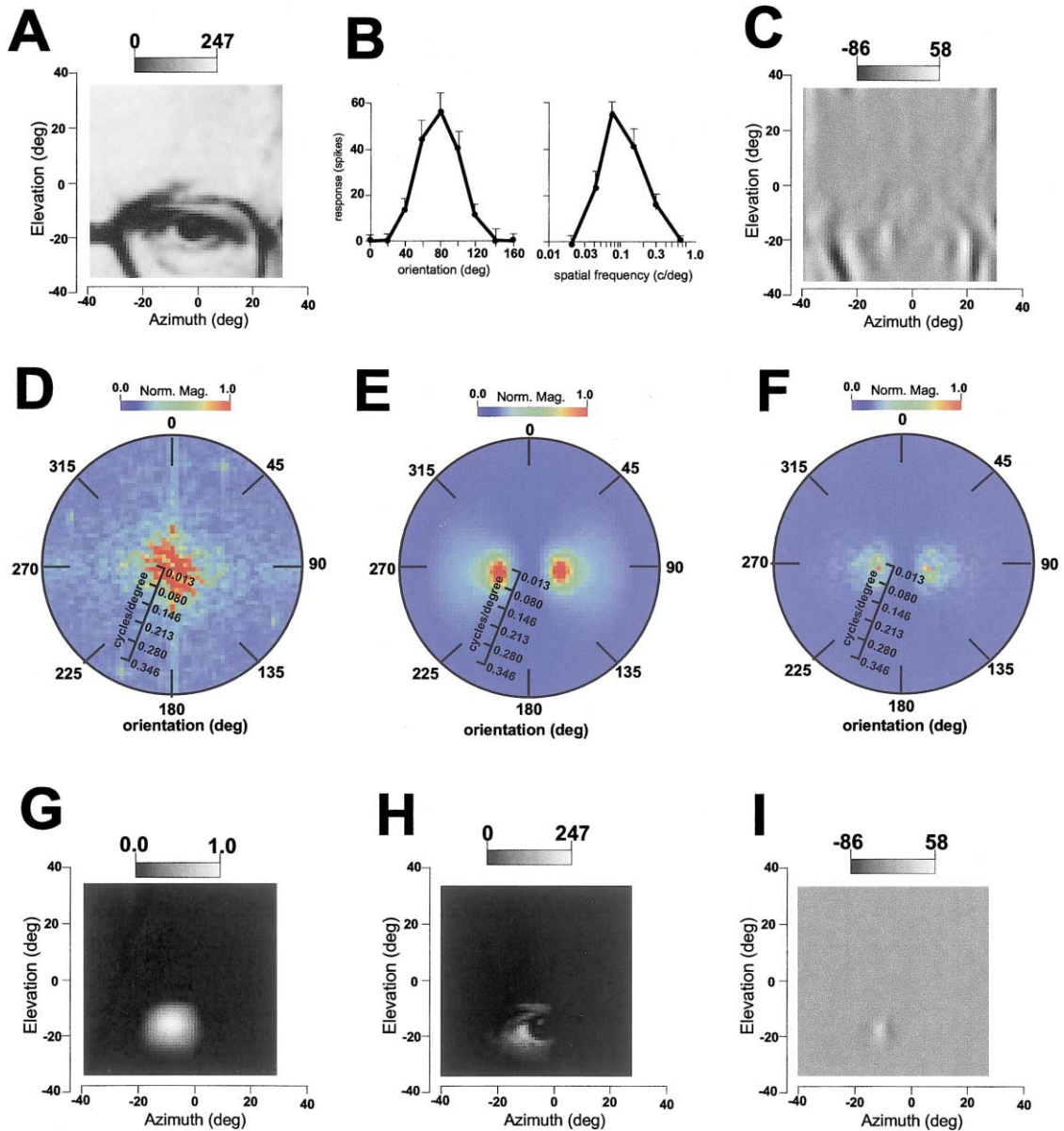
Figure 5. Method for Extracting the Orientation and Spatial-Frequency Components of a Natural Scene Falling within the Measured Classical Receptive Field at Each Recording Site

(A) The central region of a presented natural scene in the vicinity of the receptive field (see the receptive field in [G]).

(B) The orientation and spatial frequency tuning curves obtained at the recording site in response to flashed sine-wave gratings.

(C) The natural scene image shown in (A) filtered by the orientation and spatial frequency tuning properties of the recording site. Dominant vertical components lie at the peak of the orientation tuning curve shown in (B) (90° represents vertical).

(D–F) Method for filtering natural scenes using the two-dimensional fast Fourier transform (FFT). (D) The 2D spatial-frequency content of the natural scene shown in (A). The color of each pixel represents the normalized coefficient magnitude (0.0–1.0). (E) The 2D spatial-frequency representation of the orientation and spatial frequency tuning properties of the recording site derived from the tuning curves in (B). (F) The 2D spatial frequency components of the natural scene bandpass filtered by the orientation and spatial-frequency tuning properties of the recording site. This is obtained by multiplying panels (D) and (E).

(G–I) Extraction of the portion of the natural scene falling within the classical receptive field. (G) Measured receptive field (normalized from 0.0 to 1.0). (H) Extraction of the region of the original natural scene falling within the receptive field. This is obtained by multiplying panels (A) and (G). (I) Extraction of the region of the bandpass filtered natural scene falling within the receptive field. This is obtained by multiplying panels (C) and (G).

gion of the surround influencing responses within the classical receptive field.

We next investigated whether some of the variation in activity observed at individual recording sites could be accounted for by differences in the global statistical properties of each image. If this were true, then it might be expected that responses at all cortical recording sites would be modulated up or down as a group by different
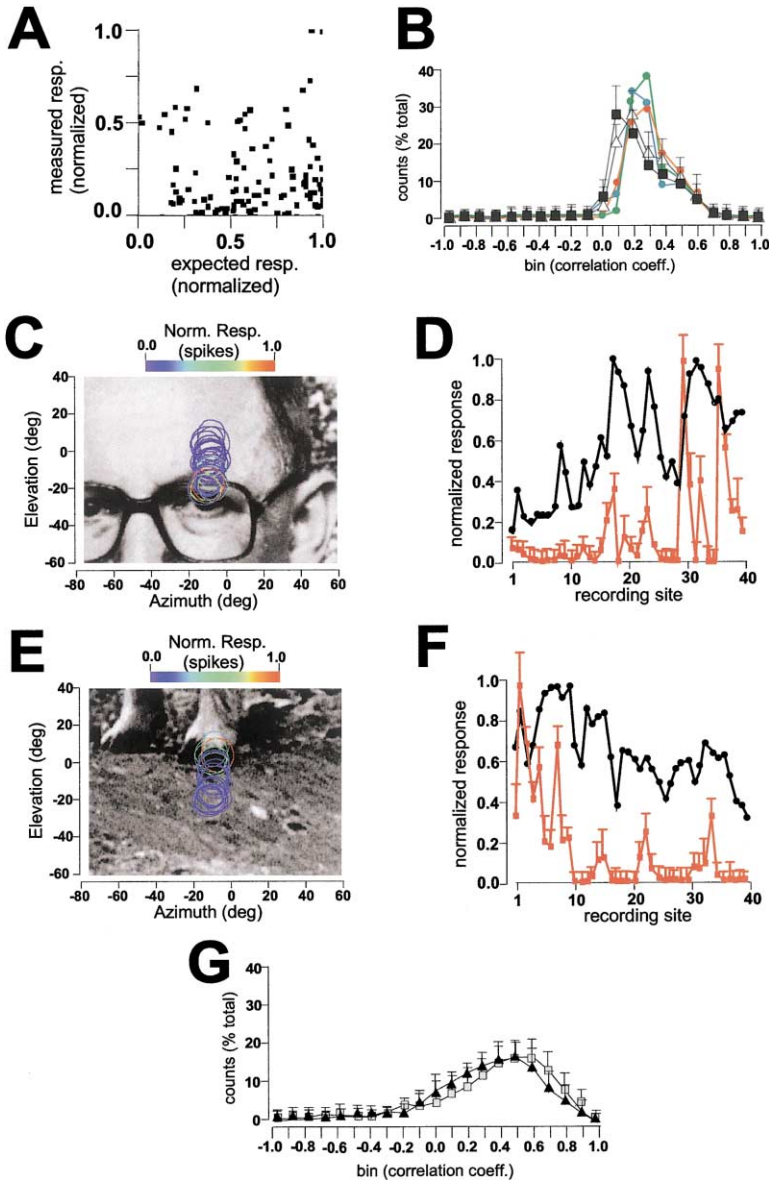
Figure 6. Correspondence of Measured and Expected Responses at Individual Sites and Across Simultaneously Recorded Sites to Natural Scenes

(A) Scatter plot of measured versus expected responses for 116 images at one recording site. The correlation coefficient (r) between the expected and measured responses is 0.03.

(B) Distribution of correlation coefficients (r) between expected and measured responses at individual recording sites to 116 natural scenes. Each r value represents the correlation between the measured and expected responses at one recording site to all images as shown in (A). The plots represent mean histograms across all animals. Black squares: distribution obtained from voltage thresholding the signal. White triangles: distribution obtained from rectified/integrated signal. Red, blue, green circles: distribution obtained from recomputed expected responses using 2-, 4-, and 6-fold diameter increases of receptive fields, respectively.

(C) Alignment of measured receptive fields obtained simultaneously at 40 recording sites in one animal to a natural scene image. Each receptive field is represented by a ring determined by thresholding the measured receptive field. The color of each receptive field represents the normalized response to the natural scene obtained at that recording site.

(D) Comparison of expected (black) and measured (red) responses obtained across the 40 recording sites shown in (C). Recording sites are ordered from 1 to 40 according to their relative elevation shown in (A), where site 1 is the topmost receptive field (i.e., highest elevation). The correlation between the expected and measured responses is $r = 0.38$. Error bars are SD for 20 presentations of the image.

(E–F) same as (C)–(D) for another natural scene. The correlation between measured and expected responses in (F) is $r = 0.21$.

(G) Distribution of correlation coefficients (r) between expected and measured responses across all simultaneously recorded sites to each of 116 natural scenes. Each r value represents the correlation between measured

and expected responses across all recording sites as shown in (D) and (F) to each image. The plots represent mean histograms across all animals. Black triangles: distribution obtained from voltage thresholding the signal. Gray squares: distribution obtained from rectified/integrated signal.

images. This hypothesis was confirmed by finding that there was a significant joint association between the mean response to each image (computed across all recording sites) and the mean image luminance and variance ($R = 0.71 \pm 0.16$, $p < 0.001$, *coefficient of multiple correlation*). Global image statistics such as mean luminance and variance were computed across the entire region of each natural scene presented to the animal, typically covering 140° by 100°. This result is consistent with previous findings that the contrast (Levitt and Lund, 1997; Polat et al., 1998) and mean illumination (Rossi et al., 1996) of surrounding stimuli can modulate the responses to stimuli placed within the classical receptive field. We also found that the computed variance of responses across all recording sites to each individual image (mean = 2.04 ± 0.63) was smaller than the com-

puted variance of responses across all recording sites to all images (mean = 5.2 ± 0.98) ($p < 0.01$, t test). Thus, the global statistical properties of different images seemed to drive the entire cortical population to a different degree.

We uncoupled the correlation measurement from these global modulating effects by computing the correlation between measured and expected population responses to each individual image (Figures 6C–6F). When computing the correlation between measured and expected population responses across all recording sites to each of the 116 natural scene images, correlation coefficients were significantly higher for the population responses than for responses at individual recording sites ($p < 0.001$, t test). For 40% of images, correlation coefficients between measured and expected popula-

tion responses were well correlated above $r = 0.6$ as compared to only 3% at individual sites (Figure 6G). Thus, after factoring out the differences in overall population activities due to global image properties, we found an increased, but still very modest, correlation between predicted and measured activity of cells. This modest correlation was not a result of random intertrial fluctuations at individual sites. This is demonstrated by the small standard deviation of responses at each recording site compared to the differences in mean responses seen between recording sites (Figures 6D and 6F). Results obtained using voltage thresholding and rectification/integration of surface recorded signals were statistically indistinguishable ($p > 0.05$, t test) (Figure 6G). Additionally, results obtained using 2- to 6-fold larger receptive fields for computing expected responses were statistically indistinguishable from results obtained using the original receptive field sizes ($p > 0.05$, t test).

Relation between Local Contrast Structure and Retinotopically Averaged Population Activity

Next, we asked whether a simple population coding model in which activity, integrated across distributed cortical sites having retinotopically overlapping receptive fields, carries the information of the local contrast structure of natural scenes. For our comparison, a method for combining neural responses into a population code and a measure of the total local contrast energy needed to be defined. For computing the population response at a given point in the visual field, we included all recording sites whose receptive fields covered that location (Figure 7A) and used the receptive field envelope to derive the weighted average of these responses (Figure 7B). This procedure was carried out at the receptive field centers of all simultaneously recorded cortical sites. This same method was applied for both deriving the measured and expected population response, with the only difference being that in the first case the measured cell responses were used as variables while in the second case the predicted model responses were used.

Based on this approach, we computed the measured and expected retinotopically integrated population response at the center pixel of each site's receptive field and derived a correlation measure for each image (Figures 7C and 7D). The degree of correlation between measured and expected retinotopic population responses was dramatically improved over the one-to-one correlation of measured and expected activity across multiple recording sites ($p < 0.001$, t test) (Figure 7E). For 81% of the images, measured and expected retinotopic population responses were well correlated, with values above $r = 0.7$ (Figure 7E). This indicates that at each point in the visual field, the integrated activity of cells whose receptive fields overlap that point is closely related to the underlying local contrast structure, as represented by the magnitude of 2D spatial frequency energy falling within these cells' classical receptive fields and bandpass tuning. Results obtained using both voltage thresholding and rectification/integration of surface recorded signals were statistically indistinguishable ($p > 0.05$, t test) (Figures 7E and 7F). Additionally, results obtained using 2- to 6-fold larger receptive fields for

computing expected responses were statistically indistinguishable from results obtained using the original receptive field sizes ($p > 0.05$, t test).

Since taking a weighted average is one form of low-pass filtering the data, it is possible that the improved correlation is due to this filtering rather than to a true correspondence between population activity and underlying contrast structure. This hypothesis was tested by recomputing the correlation between the measured and expected integrated population responses across all recording sites for each image, but now using the corresponding measured and expected mean activity of randomly chosen sites rather than of retinotopically overlapping neighbors. The number of the randomly chosen elements for each electrode was the same as the number of overlapping elements found for that electrode based on the coverage of the receptive fields. If averaging across recording sites simply removes variance, we would expect no difference between the correlations based on random elements or retinotopic neighbors. However, if the mean activity of retinotopically aligned sites is closely related to the underlying contrast structure of the image, there should be a large difference between the two correlation values. When averaging was performed across randomly chosen recording sites, correlations between measured and expected responses were significantly lower than those obtained with retinotopically aligned sites ($p < 0.01$, t test) (Figure 7E). The distribution of correlations was similar to that obtained across recording sites when no retinotopic averaging was applied ($p > 0.05$, t test) (compare Figure 6G to Figure 7E). This result confirms that cell activity, distributed across retinotopically overlapping cortical sites, is statistically linked to the local contrast structure of natural scenes. Results obtained using both voltage thresholding and rectification/integration of surface recorded signals were statistically indistinguishable ($p > 0.05$, t test) (Figures 7E and 7F).

To further explore the population code in visual cortex, we extended the computation of retinotopically averaged activity from using only the centers of the receptive fields to all pixels lying within the aggregate region of the image covered by all receptive fields (Figure 7G). The result of this extension is that we could compare a full 2D retinotopic map of measured and expected activity. High values in the expected response map indicate points in the image where many strong contrast variations could be found, whereas high values in the measured response map indicate points where the mean response of retinotopically overlapping cells was very strong. There was a high correlation between expected and measured response maps across our full image set (Figure 7E). This confirmed our finding at the centers of receptive fields that retinotopically integrated activity closely echoes the underlying local image contrast structure. To demonstrate this further, Figure 7G shows response maps superimposed upon the original image that was used to obtain the predicted and measured values. Both maps showed the highest level of activity where the contrast variations of the image were the highest. Results obtained using both voltage thresholding and rectification/integration of surface recorded signals were statistically indistinguishable ($p > 0.05$, t test) (Figures 7E and F).
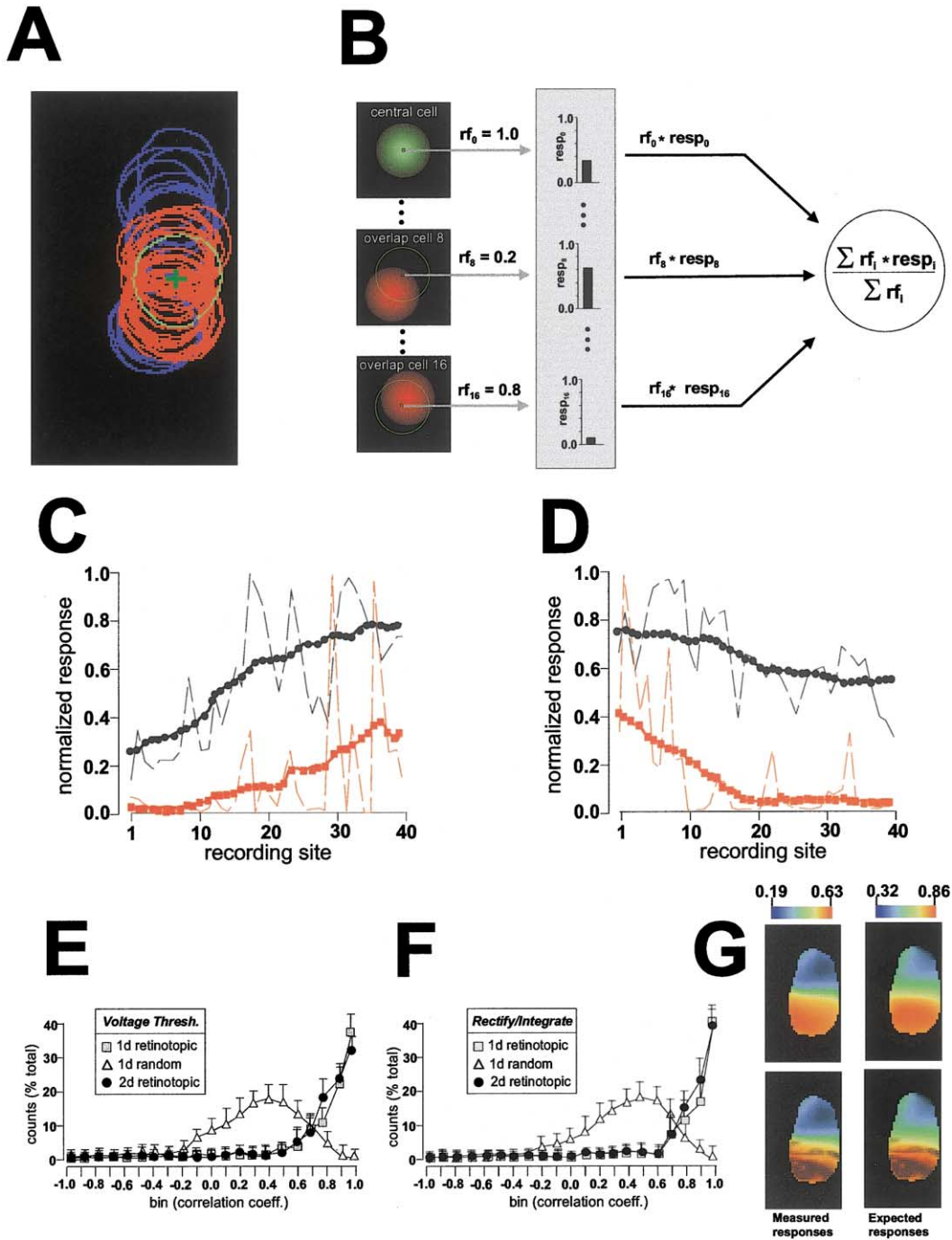
Figure 7. Integration of Responses across Cortical Sites with Retinotopically Overlapping Receptive Fields

(A) A diagram showing the receptive fields obtained simultaneously at 40 recording sites in one animal. Each receptive field is represented by a ring determined by thresholding the measured receptive field at each site. The green cross indicates the center pixel of the receptive field that is colored green. All other receptive fields that overlap the pixel marked by the green cross are colored red. The remaining receptive fields are blue.

(B) The peak amplitude of each receptive field is first normalized to 1.0. At the marked pixel, the amplitude of each receptive field that covers this point is determined (rf_i). The response at each site ($resp_i$) (graphs shown in the gray box) is multiplied by rf_i. The weighted responses are summed together and then normalized by the summation of all rf_i.

(C and D) Comparison of expected (black circles) and measured (red squares) retinotopically integrated responses computed at the receptive field center of each of 40 recording sites in one animal as described in (A) and (B). The dotted lines show the original responses at each of the 40 sites (black, expected; red, measured). (C) Integrated responses to the image shown in Figure 6C. The correlation between the expected and measured integrated responses is $r = 0.93$. (D) Integrated responses to the image shown in Figure 6E. The correlation between the expected and measured integrated responses is $r = 0.86$.

(E) Distribution of correlation coefficients between expected and measured retinotopically integrated population responses to 116 natural

The Effect of Center-Surround Interactions on the Correlation between Local Contrast Structure and Cortical Responses

All of our results so far have been obtained using large-field natural scene images. We next examined the degree to which the correlations between cortical responses and local image contrast structure depend upon center-surround interactions. If the difference between the cortical responses and the magnitude of the local contrast energy is caused by modulations due to stimulation of the receptive field surround, then restricting the stimuli to the classical receptive field should strongly improve the correlation between measured and expected values. In order to test this, surface recordings were collected in four separate animals utilizing a subset of 15 images selected from our library of 116 images. At each recording site, responses were obtained to (1) large-field natural scenes in which regions of the image fell outside the classical receptive field and (2) masked image patches from the same scenes that only covered the measured classical receptive field. The responses at each recording site for the masked condition were obtained individually.

When computed at each recording site across all images, expected responses had a similarly low correlation with measured responses to both large-field images and image patches ($p > 0.1$, t test) (Figure 8A). When computed across all recording sites for each individual image, correlations in both conditions improved significantly over correlations observed at individual sites ($p < 0.05$, t test); however, expected population responses were slightly more correlated with measured population responses to large-field images than with measured population responses to image patches ($p < 0.05$, t test) (Figure 8B). Finally, we found a similarly high correlation between expected and measured retinotopically averaged responses to both large-field images and image patches ($p > 0.1$, t test) (Figure 8C).

These results show that surround effects are not the main source of difference between local contrast energy and cell responses. If anything, they slightly improved correlations between measured and expected responses to each image when computed across multiple recording sites.

Lifetime Sparseness, Population Sparseness, and Dispersal of Cell Responses

For calculating lifetime and population sparseness, we used the formula proposed by Rolls and Tovee (1995) as modified by Willmore and Tolhurst (2001) (see Experimental Procedures). These measures of sparseness are in good agreement with other types of measurements, such as the kurtosis of the response distribution (Willmore and Tolhurst, 2001). At each recording site, a single lifetime sparseness measurement was computed across all large-field images for the measured and expected responses separately (n = 5 animals, 116 images). Measured activity had significantly higher mean lifetime sparseness than expected activity (Figure 8D) ($p < 0.001$, t test). Population sparseness was computed across all recording sites for each individual image for both expected and measured responses (Figure 8D). Similar to lifetime sparseness, measured population sparseness was significantly higher than expected population sparseness ($p < 0.001$, t test).

Dispersal was calculated for each recording site according to the method described in Willmore et al. (2000). Specifically, the variance of each recording site to all natural images was computed, normalized by the highest value, and rank ordered in a plot for the measured and the expected values separately, creating a *scree plot* (Figure 8E). The area under the scree plot quantifies the distribution of variance across the recording sites and also indicates the dispersal of cell responses across the set of stimuli. A flat plot having high dispersal means that the cells contributed equally to coding of the stimulus set. Mean dispersal of the expected (135.1 ± 11.6) and measured (126.5 ± 15.9) responses was not statistically different ($p > 0.1$, t test).

In our separate group of animals (n = 4 animals, 15 images), each calculation of lifetime sparseness, population sparseness, and dispersal was recalculated for image patches presented in the classical receptive field, and large-field images that stimulated wide areas beyond the classical receptive field. There was no change in lifetime sparseness for patches and large-field images ($p > 0.05$, t test) or in population sparseness for patches and large-field images ($p > 0.05$, t test) (Figure 8F). Mean dispersal of responses to patches (102.3 ± 24.1) and large-field (121.2 ± 16.3) images was statistically indistinguishable ($p > 0.1$, t test)

Discussion

Neural activity was simultaneously recorded across extended areas of primary visual cortex in response to natural scene images. We assessed the extent to which this activity conformed to expected activity based on the classical model of cell responses to sine wave grating stimuli. Our results demonstrate several points that will be discussed as follows: (1) the relation between surface and penetrating electrode recording, (2) the relation

---

scenes. Data are obtained from voltage thresholding the surface recorded signal. Gray boxes show the results of retinotopically integrated responses at the center pixel of each recording site's receptive field as described in (A) and (B). White triangles show the results of responses averaged across each recording site and randomly selected sites. Black circles show the results obtained from 2D maps of retinotopically integrated activity shown in (G). The distribution of correlation coefficients for 1D and 2D retinotopically integrated activity are not statistically different ($p > 0.15$, t test), but are both significantly higher than for randomly integrated activity ($p < 0.001$, t test). Each plot represents the mean histogram across all animals.

(F) Same as (E), but data are obtained from a rectified/integrated surface recorded signal.

(G) Measured (left column) and expected (right column) 2D maps of retinotopically integrated activity using the same method described in (A) and (B), except the computation is performed at all pixels lying within the aggregate region covered by all receptive fields. The correlation between the expected and measured response maps is $r = 0.95$. Color encodes the integrated responses at each pixel. In the bottom row, measured and expected 2D maps of integrated activity are overlaid upon the aligned stimulus image.
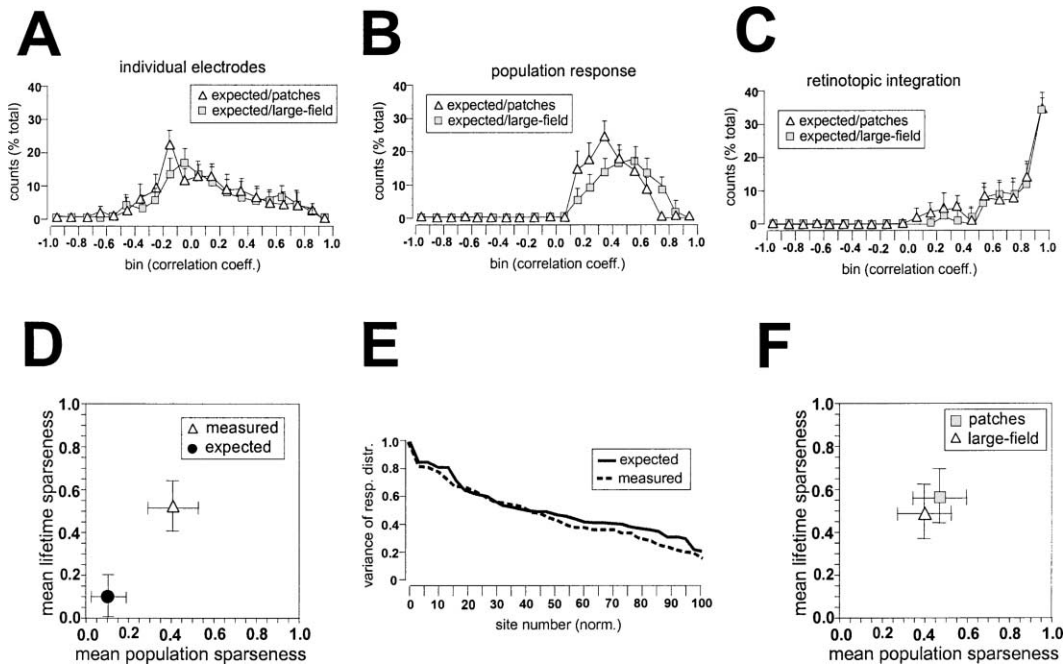
Figure 8. Analyses of Center-Surround Interactions and Sparseness

(A–C) Analysis of center-surround interactions across all animals. (A) Histogram of *r* (correlation between expected and measured responses) computed at individual sites across 15 images. White triangles show the correlation between expected responses and measured responses to masked image patches. Gray squares show correlation between expected responses and measured responses to large-field images. (B) Histogram of *r* computed across simultaneously recorded cortical sites for each image. Legend is the same as for (A). (C) Histogram of *r* computed across retinotopically integrated activity at the centers of all sites' receptive fields for each image. Legend is the same as for (A).

(D) Plot of mean population sparseness versus mean lifetime sparseness for measured and expected responses to large-field natural scenes. Mean values are computed across all animals. Error bars are SD.

(E) Dispersal of measured and expected responses to large-field natural scenes. Plots show mean dispersal averaged across all animals. The number of recording sites obtained in each animal is normalized to 100 to allow averaging across animals with varying numbers of recording sites.

(F) Plot of mean population sparseness versus mean lifetime sparseness for responses to image patches and large-field natural scenes. Mean values are computed across all animals. Error bars are SD.

among local contrast structure, single cell responses, and population coding, and (3) the relation between sparseness and efficient coding.

Surface versus Penetrating Electrode Recording

The cortical surface recording technique we have described provides a robust method for high-density mapping of neuronal activity across exposed cortical regions. This technique allows the recording of spatially localized cortical activity, with response properties similar to a layer 2/3 recording, without having to penetrate the brain. The method offers a distinct advantage over the use of penetrating electrodes for mapping brain activity at high spatial and temporal resolution, since the use of large numbers of penetrating electrodes will likely cause significant damage to blood vessels and cortical tissue. With some modifications, the technique also has potential as a chronic recording method that is less traumatic than traditional penetrating electrode techniques.

Using the standard description of receptive fields, orientation, spatial frequency, and contrast tuning, we found no significant differences between recorded activity on the surface compared to activity recorded with penetrating electrodes in layer 2/3. This might not be

considered too surprising given the fact that the estimated area from which penetrating electrodes collect multi-unit spike activity is about 200 μm (Logothetis et al., 2001), about the same distance as our layer 2/3 recording from the surface. Although, signal collected from the surface might also be expected to be very noisy since it would possibly include action potentials arising from axons of passage or dendrites that ramify in the vicinity of the recording site, we did not find this to be true.

We propose that one reason for the high correlation between results with surface and penetrating recording is due to the fact that the measures we used to describe cell behavior (e.g., orientation tuning, spatial frequency tuning, receptive field size, responses to natural scenes) are overall characterizations of cell activity in the superficial layers. Our results demonstrate that surface recording has no noticeable effect on the quality of measurements of these characteristics of local cell populations when compared to penetrating recordings.

Local Contrast Structure, Single-Cell Coding, and Population Coding

We found no evidence that cell activity at individual recording sites directly encodes local contrast informa-

tion of natural scene stimuli even when we compensated for the variations due to global image characteristics. There are several earlier studies pointing out that simple oriented spatial frequency filter models do not describe adequately cell behavior in the striate cortex. However, these reports focused either on the nonlinearity of the complex cells due to their phase invariance (Skottun et al., 1991), or center-surround interactions (Allman et al., 1985; Walker et al., 2000). Our recordings were dominated by phase-insensitive responses, and therefore, it was appropriate to use a contrast energy model which has been used by other studies to model complex cell responses (Lades et al., 1993; Willmore et al., 2000). Thus, this could not be the source of the difference between the expected and measured values. We also showed that surround effects do not contribute significantly to our results.

Although the correlation between local contrast structure and cell responses is modest at the level of individual cortical sites, a very simple population code, derived from activity integrated across cortical sites having retinotopically overlapping receptive fields, represents the local contrast structure of natural scenes very well. The significant difference obtained by averaging across retinotopically neighboring sites as opposed to randomly chosen sites (see Figure 7E) excludes the possibility that simple averaging or lowpass filtering is responsible for this improvement. Additionally, because the standard deviation of responses at each recording site is small compared to the differences in mean responses between recording sites (see Figures 6D and 6F), this further excludes the possibility that the activity at individual sites is inherently noisy or unreliable and that averaging simply helps to restore the signal. While we do not yet know how to precisely interpret this cortical code, our results indicate that it is more complex than a linear coding of the local contrast energy within the classical receptive field of individual cells. However, our results demonstrate that by integrating across retinotopically neighboring recording sites, a significant degree of linearity is restored to the distributed representation of natural scenes in primary visual cortex.

The Gabor filter model of V1 processing, which generates a multiscale representation of local spatial frequency and orientation components falling within cells' classical receptive fields, has been challenged on many levels. In this light, our study is a restoration of this original classical model claiming that relevant information for coding natural scenes is in the classical receptive field. This notion is bolstered by our finding that surround effects are not the main source of difference between local contrast energy and cell responses, but rather these differences are primarily produced by stimuli directly within the classical receptive field. Thus, the main point of our departure from the classic view is that our results suggest that the local contrast energy in natural scenes is coded by population activity distributed across retinotopically overlapping cortical sites, not at individual cortical locations. The high correlation between the weighted average response of retinotopically overlapping recording sites and the local contrast structure of natural scenes means that any cell in higher visual cortical areas can extract this contrast information in a biologically plausible way if necessary.

Sparseness and Efficient Coding in the Striate Cortex

High sparseness of cell responses in the brain has been proposed to be advantageous on two different grounds. First, because minimizing the number of active cells at any moment reduces the metabolic cost of functioning (Field, 1994), and second, because it improves the coding efficiency of the brain (Barlow, 1989; Treves and Rolls, 1991). The metabolic cost can be fully characterized by lifetime sparseness, which describes the individual cell responses to an appropriate set of stimuli by using some simplifying assumptions. However, the assessment of coding efficiency, as used by Treves and Rolls (1991), is based upon the assumption that different stimuli activate as different a subset of cell responses as possible, while similar inputs activate more similar subsets. These requirements of efficient coding are based upon the instantaneous population activity to stimuli and cannot be guaranteed by high lifetime sparseness itself. The assessment of coding efficiency requires the use of all three lifetime sparseness, population sparseness, and dispersal measurements, and it needs to measure dispersal carefully so that spurious cell responses could not distort the shape of the scree plot (Willmore and Tolhurst, 2001). High dispersal together with high population and high lifetime sparseness is a good indicator of a sparse-distributed representation that, in turn, suggests both metabolic efficiency and efficient information coding for pattern storage and associative learning.

Our method of assessing the efficiency of coding in the striate cortex benefits from using all three of the aforementioned measurements. Furthermore, because our model generated the expected cell responses based on real receptive field sizes, contrast functions, spatial frequency, and orientation preferences, it also avoids the problems originating from arbitrary assumptions about these parameters. We emphasize that our measurements of sparseness and dispersal are based on multi-unit signals, whereas a precise assessment of these quantities would require clearly isolated single-unit measurements. In addition, there is no guarantee that our ad hoc set of natural images is unbiased in their correlations and they span the entire input space relevant for the cells. Nevertheless, we are able to derive meaningful conclusions from our data for two reasons. First, collecting spike responses indiscriminately from multiple cells will have the effect of reducing sparseness on the average, creating a lower bound for the true sparseness values. Thus, while a low sparseness value would not guarantee that the participating cells indeed have low sparseness, a high sparseness of multi-unit responses is a good indicator of at least as high individual sparseness on average. Second, instead of emphasizing absolute values, our conclusions are based on relative measures since we are comparing our recorded values to those expected from our linear filter model. Thus, we expect that biases introduced by our data set affect the recorded and expected values equally.

Based on our measurements, we found that cells in the primary visual cortex of the ferret are functioning highly efficiently. The mean lifetime sparseness of measured responses to natural scenes was 0.51 compared to 0.12 for the mean sparseness of expected values,

and we found the same pattern with population sparseness with an only slightly lower mean (0.42). Computing the measured mean lifetime sparseness according to the formula used by Vinje and Gallant (2000), we obtained 51%, which is slightly higher than the value they found stimulating the classical receptive field of cells in the striate cortex of the awake monkey (41%), but slightly lower than the value they reported when stimulating an area four times the size of the classical receptive field (62%). In contrast to Vinje and Gallant (2000), we found no significant change in mean lifetime sparseness when the stimulation was restricted to the classical receptive field as compared to large-field stimulation. This might indicate either significant differences in center-surround interaction mechanisms between ferrets and monkeys, or significant differences between cell responses in awake animals and anaesthetized preparations.

The mean areas under the scree plots that measure the dispersal of the codes (referred to as dispersal value) were 135.1 and 126.5 for the expected and measured response variances (see Figure 8E). The expected dispersal value is significantly higher than the value reported in Willmore et al. (2000) for a Gabor filter coding model. This is likely due to the same reason our sparseness values deviated from theirs in the opposite direction: in our code there were no filters with substantially lower spatial frequency preference that could encode the majority of the variance in the images. Our dispersal values are as high as the values reported by Willmore et al. (2000) for the Olshausen and Field (1996) sparse code, the most successful coding scheme of all tested schemes from the standpoint of dispersal measures.

Experimental Procedures

Surface Electrode Recording and Signal Conditioning
Electrodes were made from 0.025 mm diameter insulated tungsten wires (California Fine Wire), which contacted the pia. The electrode design incorporated a 1 cm long flexible horizontal support arm that allowed the recording tip to move up and down under light force. A 0.0125 mm tungsten wire was used as a differential electrode. Wires were connected to 18 inch shielded miniature coaxial cables that terminated in a 16 channel amplifier module. Each row of 15 electrodes plus differential was connected to a separate amplifier module. Amplifiers provided a gain of 10,000. Two-stage RC circuits bandpass filtered the signal between 600 and 6000 Hz. The voltage output from each electrode amplifier was digitized at a 10 kHz/channel sampling rate using a 64 channel data acquisition card (National Instruments) running under LABVIEW (National Instruments) and Windows NT.

Two methods were used for quantifying the neural activity signal obtained from surface recordings. In the first method, multi-unit spike waveforms were extracted from the digitized raw signal by measuring the standard deviation of the noise on each channel separately and then setting a voltage detection threshold at 4.0 to 4.5 times this value. In the second method, the absolute value of the recorded signal was obtained, and the resulting full-wave rectified signal was integrated. All results described in the paper were obtained using voltage-thesholded surface recorded signals unless otherwise specified.

Excess fluid was dried from the cortical surface and checked every 15 to 30 min for buildup. If necessary, the cortical surface was sprayed with saline to keep it moist. Stable recordings were maintained for up to 14 hr.

Animal Preparation
All experimental procedures were approved by the University of Rochester Animal Care and Use Committee. Anesthetized, paralyzed adult male sable ferrets were used for electrophysiological recording. See Weliky and Katz (1997) for animal preparation methods. A hole (10 mm by 5 mm) was made in the cranium over area 17, and the dura was removed. Each optic disk was plotted on the display screen with a reversing ophthalmoscope. The vertical meridian was defined as the midpoint between the projections of the two optic disks. Visual stimulation was presented to the contralateral eye. The electrode array was placed mediolaterally along the caudal pole of the lateral gyrus, corresponding to the location of area 17. Data from cortical sites that exhibited (1) increased receptive field size or (2) a reversal of receptive field position along the horizontal axis were not used for analysis, since these sites were likely located in area 18.

Stimulus Presentation and Data Acquisition
Stimuli were generated on a Macintosh G4 and displayed by an OPTIMA EzPRO 710 XGA video projector at a refresh rate of 75 Hz. Stimuli were projected onto a 4 by 3 foot rear projection screen placed 48 cm in front of the animal with a mean luminance of 30 cd/m². The output of the display projector was γ corrected for linearity. All stimuli, including gratings, squares, and images, were flashed onto the display screen. Data were acquired during a 750 ms stimulus presentation trial as follows. After a 250 ms baseline, the stimulus was flashed on the screen for 250 ms, and the recording was continued for another 250 ms after stimulus offset. For voltage-thresholded signals, the number of spikes acquired during the ON and OFF portion of the stimulus were summed together and subtracted from the number of spikes, multiplied by two, obtained during the 250 ms baseline period. For rectified/integrated signals, the integrated signal obtained during the ON and OFF portion of the stimulus was subtracted from the integrated signal, multiplied by two, obtained during the baseline period.

Receptive fields at each cortical recording site were simultaneously mapped with a two-dimensional grid of high contrast squares that covered the aggregate visual area of all receptive fields. Squares were approximately 2° on a side, and the two dimensional grid was typically 10 to 20 squares in length along each side. During a single trial, each square was individually flashed in random order. Five to eight trials were averaged to construct a receptive field plot. The orientation and spatial frequency tuning, as well as contrast response, were simultaneously obtained at all recording sites using flashed sine wave gratings. The spatial coverage of the sine wave gratings typically covered an area 15%–25% larger than the aggregate visual area of all receptive fields. Nine different orientations in 20° increments, four contrasts in 0.25 increments from 0.25 to 1.0, and six spatial frequencies were used. The spatial frequencies were determined during each experiment to span the minimum to maximum spatial frequency tuning of all recorded sites. During a single trial, each grating was presented in random order at two phases shifted by 90°. All combinations of orientations, spatial frequencies, phases, and contrasts were presented for a total of 432 stimuli. Six to ten trials were averaged.

Due to time constraints, it was impractical to present the full stimulus set at each recording site restricted to the classical receptive field. However, the following procedure was carried out that closely provided this information. Previous studies have demonstrated that cell responses are suppressed (Walker et al., 2000), uniformly without altering the overall shape or peak of the orientation tuning curve (Sillito et al., 1995), as a grating is incrementally expanded beyond the classical receptive field. Therefore, we determined the peak response at each site to the optimal high contrast orientation/spatial frequency grating masked to the receptive field. Six to eight trials were averaged for the masked optimal grating presented at each site. We then scaled the orientation tuning curves obtained from the full set of 432 grating stimuli (covering the aggregate visual area of all RFs) by the peak response to the optimal high-contrast orientation/spatial frequency grating masked to the receptive field, divided by the peak response to the same optimal orientation/spatial frequency grating that covered the aggregate visual area of all receptive fields. This produced at each site a set of orientation tuning curves that were scaled to the appropriate response amplitude if the gratings had been masked to the classical receptive field.

116 natural scene images including people, animals, and landscapes were randomly presented during each trial. Data were collected and averaged for each image from twenty trials. Presentation trials of gratings, and images were interleaved throughout the recording session. Receptive fields were also mapped before each trial throughout the experiment. If receptive fields were found to have drifted more than one mapping square in distance since the start of the experiment, the recording session and data collection were stopped.

At each site, responses to gratings and natural scenes were normalized by dividing by the peak response to the optimal high contrast orientation/spatial frequency grating masked to the receptive field for that particular site. This normalized the set of orientation tuning curves from 0.0 to 1.0 at each site. The reason for this normalization was due to stimulus-independent variations in spike counts that were observed between different sites. Such variability between sites appeared to arise primarily from differences in the number of recorded units, or variability of the contralateral eye in driving units due to shifts in ocular dominance between sites. This procedure removed any differences in spike activity that were introduced by stimulus-independent factors through the recording process and normalized all responses to the peak grating response at each site.

Construction of the Receptive Field Model

The 2D frequency domain filter was constructed from orientation and spatial frequency tuning curves, normalized from 0.0 to 1.0, and obtained at 1.0 contrast. The 2D fast Fourier transform (FFT) was computed of each image, and this result was multiplied by the 2D frequency domain filter. Following this multiplication, the inverse 2D FFT was computed. The patch of the resulting filtered image that fell within the classical receptive field was extracted by multiplying this image by the mapped receptive field normalized from 0.0 to 1.0. Image patches contained both positive and negative values as a result of the Fourier transform. In order to compute the Michelson contrast of these patches, the image pixel values had to be shifted to all positive values (this procedure was unnecessary when computing the RMS contrast). The minimum negative pixel value in the entire set of filtered image patches was subtracted from all of the images. This shifted upwards the pixel values in all filtered image patches in the set by an equal amount such that the minimum pixel value in the set was 0.0. The Michelson contrast of these filtered image patches was computed according to [contrast = (max − min)/(max + min)] and then multiplied by the measured nonlinear contrast response function. The contrast response function was obtained from the experimentally measured peak firing rates at the preferred orientation/spatial frequency at each contrast, normalized from 0.0 to 1.0. All analysis software was written in LABVIEW.

Center-Surround Interactions

Stimuli were presented as previously described. Data were collected and averaged for each full-field image and masked image patch from 10 trials. Trials consisting of large-field images and masked image patches were randomly interleaved throughout the experiment. Receptive field plots and the assessment of receptive tuning properties were performed as previously described. Each image patch was presented within a circular mask the size of the measured receptive field. To prevent boundary effects, the radius of each circular mask was increased by 10%, and pixels within this outer 10% annulus were linearly blended into a neutral gray background.

Sparseness Measurement

Both lifetime and population sparseness measures were computed according to the following formula:

$$S = 1 - [\Sigma_{i=1,N} (r_i/N)]^2/[\Sigma_{i=1,N} (r_i^2/N)].$$

For lifetime sparseness, S is the lifetime sparseness at a given recording site, N is the number of presented images, and r_i stands for the evoked response to the ith stimulus. For population sparseness, S is the population sparseness of a given image, N is the number of recording sites, and r_i stands for the response of the ith recording site when presented with the image. For both measures S changes between 0 and 1 where larger values indicate high sparseness with only a few images strongly activating a given recording

site (lifetime) or only a few active recording sites for any specific image (population).

References

Albrecht, D.G., and Hamilton, D.B. (1982). Striate cortex of monkey and man: contrast response function. J. Neurophysiol. 48, 217–237.

Allman, J.M., Miezin, F., and McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neuropsychological mechanisms for local-global comparisons in visual neurons. Annu. Rev. Neurosci. 8, 407–430.

Barlow, H.B. (1989). Unsupervised learning. Neural Comput. 1, 295–311.

Blakemore, C., and Tobin, E.A. (1972). Lateral inhibition between orientation detectors in the cat's visual cortex. Exp. Brain Res. 15, 439–440.

Bringuier, V., Chavane, F., Glaeser, L., and Fregnac, Y. (1999). Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. Science 283, 695–699.

Chao-Yi, L., and Wu, L. (1994). Extensive integration field beyond the classical receptive field of cat's striate cortical neurons-classification and tuning properties. Vision Res. 34, 2337–2355.

DeAngelis, G.C., Cumming, B.G., and Newsome, W.T. (1998). Cortical area MT and the perception of stereoscopic depth. Nature 394, 677–680.

DeAngelis, G.C., Robson, J.G., Ohzawa, I., and Freeman, R.D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. J. Neurophysiol. 68, 144–163.

DeValois, K.K., and Tootell, R.B.H. (1983). Spatial-frequency-specific inhibition in cat striate cortex cells. J. Physiol. 336, 359–376.

DeValois, R.L., Albrecht, D.G., and Thorell, L.G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. Vision Res. 22, 545–559.

Field, D. (1994). What is the goal of sensory coding? Neural Comput. 6, 559–601.

Fries, P., Roelfsema, P.R., Engel, A.K., Konig, P., and Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. Proc. Natl. Acad. Sci. USA 94, 12699–12704.

Gabor, D. (1946). Theory of communication. Journal of IEE 93, 429–457.

Gilbert, C.D., and Wiesel, T.N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. Vision Res. 30, 1689–1701.

Hubel, D.H., and Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. J. Physiol. 148, 574–591.

Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R.P., and Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. IEEE Transactions on Computers 42, 300–311.

Levitt, J.B., and Lund, J.S. (1997). Contrast dependence of contextual effects in primate visual cortex. Nature 387, 73–76.

Logothetis, M.K., Pauls, J., Augath, M., Trinath, T., and Deltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150–157.

Nelson, J.J., and Frost, B.J. (1978). Orientation-selective inhibition from beyond the classic visual receptive field. Brain Res. 139, 359–376.

Olshausen, B.A., and Field, D.J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature *381*, 607–609.

Polat, U., Mizobe, K., Pettet, M.W., Kasamatsu, T., and Norcia, A.M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. Nature *391*, 580–584.

Rolls, E.T., and Tovee, M.J. (1995). Sparseness of the neural representation of stimuli in the primate temporal visual cortex. J. Neurophysiol. *73*, 713–726.

Rossi, A.F., Rittenhouse, C.D., and Paradiso, M.A. (1996). The representation of brightness in primary visual cortex. Science *273*, 1104–1107.

Sillito, A.M., Grieve, K.L., Jones, H.E., Cudeiro, J., and Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. Nature *378*, 492–496.

Skottun, B.C., DeValois, R.L., Grosof, D.H., Movshon, J.A., Albrecht, D.G., and Bonds, A.B. (1991). Classifying simple and complex cells on the basis of response modulation. Vision Res. *31*, 1079–1086.

Treves, A., and Rolls, E.T. (1991). What determines the capacity of autoassociative memories in the brain? Network: Computation in Neural Systems *2*, 371–397.

Vinje, W.E., and Gallant, J.L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. Science *287*, 1273–1276.

Walker, G.A., Ohzawa, I., and Freeman, R.D. (1999). Asymmetric suppression outside the classical receptive field of the visual cortex. J. Neurosci. *19*, 10536–10553.

Walker, G.A., Ohzawa, I., and Freeman, R.D. (2000). Suppression outside the classical cortical receptive field. Vis. Neurosci. *17*, 369–379.

Weliky, M., and Katz, L.C. (1997). Disruption of orientation tuning in visual cortex by artificially correlated neuronal activity. Nature *386*, 680–685.

White, L.E., Bosking, W.H., Williams, S.M., and Fitzpatrick, D. (1999). Maps of central visual space in ferret V1 and V2 lack matching inputs from the two eyes. J. Neurosci. *19*, 7089–7099.

Willmore, B., and Tolhurst, D.J. (2001). Characterizing the sparseness of neural codes. Network: Computation in Neural Systems *12*, 255–270.

Willmore, B., Watters, P.A., and Tolhurst, D.J. (2000). A comparison of natural-image-based models of simple-cell coding. Perception *29*, 1017–1040.