Isolation of Relevant Visual Features from Random Stimuli for Cortical Complex Cells

Jon Touryan,[1] **Brian Lau,**[2] **and Yang Dan**[1,2]

[1]*Group in Vision Science, University of California, Berkeley, California 94720, and* [2]*Division of Neurobiology, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3200*

A crucial step in understanding the function of a neural circuit in visual processing is to know what stimulus features are represented in the spiking activity of the neurons. For neurons with complex, nonlinear response properties, characterization of feature representation requires measurement of their responses to a large ensemble of visual stimuli and an analysis technique that allows identification of relevant features in the stimuli. In the present study, we recorded the responses of complex cells in the primary visual cortex of the cat to spatiotemporal random-bar stimuli and applied spike-triggered correlation analysis of the stimulus ensemble. For each complex cell, we were able to isolate a small number of relevant features from a large number of null features in the random-bar stimuli. Using these features as visual stimuli, we found that each relevant feature excited the neuron effectively in isolation and contributed to the response additively when combined with other features. In contrast, the null features evoked little or no response in isolation and divisively suppressed the responses to relevant features. Thus, for each cortical complex cell, visual inputs can be decomposed into two distinct types of features (relevant and null), and additive and divisive interactions between these features may constitute the basic operations in visual cortical processing.

Key words: complex cell; primary visual cortex; Wiener kernel; principal component analysis; spatiotemporal; nonlinear

An important goal in studying the receptive-field properties of visual neurons is to understand how they respond to complex spatiotemporal inputs, including those encountered in natural scenes. To analyze the responses to complex stimuli, a useful approach is to decompose the stimuli into a set of basic features (basis set) and to characterize how each feature contributes to the neuronal response. Several methods have been used to define basis sets for the efficient representation of visual stimuli, including principal component analysis (PCA), independent component analysis (Bell and Sejnowski, 1997; van Hateren and Ruderman, 1998), and/or analysis based on sparse coding (Olshausen and Field, 1996). For studying the response properties of a given visual neuron, it is desirable to construct a basis set so that the neuron responds to only a small number of visual features in the set. The segregation between a small number of "relevant" visual features and a large number of "irrelevant" features can greatly facilitate experimental characterization of the visual neuron.

For neurons with a linear stimulus–response relationship, relevant visual features can be identified by estimating their linear receptive fields using a spike-triggered average of the stimulus ensemble (also called "reverse correlation") (de Boer and Kuyper, 1968). This method has been widely used to measure the spatiotemporal receptive fields of neurons in the early visual pathway (Jones and Palmer, 1987; Reid et al., 1997); the resulting receptive fields can largely account for the neuronal responses to complex spatiotemporal stimuli (Brodie et al., 1978; Dan et al., 1996). However, in the visual cortex most of the neurons are complex cells with nonlinear stimulus–response relationships that cannot be characterized with the spike-triggered average. In the present study, we have used spike-triggered correlation analysis of the stimulus ensemble (de Ruyter van Steveninck and Bialek, 1988; Yamada and Lewis, 1999; Brenner et al., 2000) to construct the basis set for each complex cell. We found that visual features in such a set are clearly segregated into two categories: a small number of relevant features and a large number of null features. Using visual stimuli consisting of either a single feature or a combination of features, we directly measured the contribution of each type of feature to the cortical responses.

MATERIALS AND METHODS

Physiological preparation. Adult cats (2–3 kg) were initially anesthetized with isoflurane (3%, with O_2) followed by sodium pentothal (10 mg/kg, i.v., supplemented as needed). During recording, anesthesia was maintained with sodium pentothal (3 mg · kg^{-1} · hr^{-1}, i.v.), and paralysis was maintained with pancuronium bromide (0.1–0.2 mg · kg^{-1} · hr^{-1}, i.v.). The pupils were dilated with 1% atropine sulfate, nictitating membranes were retracted with 2.5% phenylephrine hydrochloride, and the eyes were mechanically stabilized and optimally refracted. End-expiratory CO_2 was maintained at 4%, the core body temperature was kept at 38°C, and the electrocardiogram and EEG were monitored continuously. All experimental procedures were performed as approved by the Animal Care and Use Committee at the University of California, Berkeley.

Recording. Extracellular recordings were made with tungsten electrodes (A-M Systems, Carlsborg, WA). Unit isolation was based on the cluster analysis of waveforms and the presence of a refractory period determined from the autocorrelograms. Cells were classified as simple if their receptive fields had clear on and off subregions (Hubel and Wiesel, 1962) and if the ratio of the first harmonic to the DC component of the response to an optimally oriented drifting grating was >1 (Skottun et al., 1991). All other cells were classified as complex. Among the 61 complex cells recorded, one was excluded from analysis because of its low firing rate in response to random-bar stimuli (<1 spike per second).

Visual stimulation. Visual stimuli were generated with a personal

computer and presented with a Barco monitor (size, 40×30 cm; refresh rate, 120 Hz; maximum luminance, 80 cd/m^2). Luminance nonlinearities were corrected using software written in our laboratory. The random-bar stimuli were presented in a rectangular patch covering the receptive field of each cell. This patch was divided into 16 bars aligned to the optimal orientation of the cell; the length of the bars was equal to or slightly longer than the receptive field. The contrast of each bar was temporally modulated according to a pseudorandom binary m-sequence (Sutter, 1987) (luminance, ± 39 cd/m^2 from the mean of 40 cd/m^2). The full m-sequence was 32,767 frames long and was updated every other frame, for an effective frame rate of 60 Hz. To measure the contrast–response functions of individual features (see Fig. 5), we randomly interleaved short movies (16 frames per movie) of relevant and null features, each at a range of contrasts (positive and negative, see below for definition of contrast), with no gap between movies. To measure the interaction between two relevant features (see Fig. 6), we generated a set of short movies containing all possible linear combinations of the two features. The number of repetitions for each short movie varied between 1 and 120, which was proportional to the probability of the corresponding feature contrast in the random-bar stimuli (for example, the probability of a high contrast for a given feature in the random-bar stimuli is generally lower than the probability of a low contrast for the same feature; thus, the movie of the feature at the high contrast was repeated fewer times). Note that in each movie, which contains either a single feature or a combination of features, the luminance of each bar must be between -1 and 1 (corresponding to 0 and 80 cd/m^2, respectively), which limits the maximum contrast of each feature that can be presented (the definition of contrast is described below).

Spike-triggered correlation analysis. In general, if certain features in the visual stimuli affect the firing probability of the cell, the spike-triggered stimulus ensemble should exhibit a different probability distribution from the entire stimulus ensemble (see Fig. 1B; compare the distribution of the *filled circles* and the distribution of all of the *circles*). Although a change in the probability distribution can be reflected in a change in the first-order (mean), second-order (variance), or higher-order moments, the correlation analysis aims to identify features with changed variance. Because PCA results in a set of components with their variance ranking from the highest to the lowest, it is ideally suited for the identification of features with outstanding variance. Practically, identification of relevant features was achieved by finding eigenvalues of the spike-triggered correlation matrix that were significantly different from the eigenvalues of the control correlation matrix (computed by randomly sampling the entire stimulus ensemble). For each cell, responses to three to four repeats of the random-bar stimuli (~ 9 min) were used for spike-triggered correlation analysis. Each pattern in the stimulus ensemble consisted of luminance at 16 bar positions at 16 frames (assuming that neuronal spiking probability depends only on the immediate stimulus history within 16 frames, lasting for 268 msec), which was uniquely specified by 256 parameters. The spike-triggered correlation matrix, $[C_{m,n}]$ ($m, n = 1, 2, \ldots, 256$) was computed as follows:

$$C_{m,n} = \frac{1}{N} \sum_{i=1}^{N} S_m(i) S_n(i),$$

where $S_m(i)$ and $S_n(i)$ are the mth and nth parameters of the stimulus pattern preceding the ith spike, respectively, and N is the total number of spikes in the response. The resulting matrix is closely related to the second-order Wiener kernel (Wiener, 1958; Marmerlis and Marmeralis, 1978) of the neuron. Eigenvalues and eigenvectors of this spike-triggered correlation matrix were then computed. To compute each control correlation matrix, we generated a random spike train with the same number of spikes as in the recorded response but with random spike timing; the correlation matrix was computed based on this simulated random spike train. Because subsequent experiments required fast identification of the significant eigenvectors, we computed only five control correlation matrices for each cell during the experiments; the confidence interval for the control eigenvalues was set at mean ± 5.2 SD (corresponding to $p < 10^{-4}$). Eigenvectors with eigenvalues outside of the control confidence interval were considered significant. Subsequent offline analyses with 100 control matrices confirmed that the significant eigenvectors were identified reliably using only five control matrices.

An important question is how the constraints of the above method affect the outcome of the analysis. For example, the eigenvectors must be orthogonal, which could affect the visual features identified. As shown in

Figure 4, the two significant eigenvectors for most complex cells exhibited similar spatial frequencies; one might suspect that the $\sim 90°$ spatial phase difference between them resulted from the orthogonality between the two vectors. However, this is not the case. First, both vectors are spatiotemporal patterns. The fact that the dot product of them is 0 (summed over all temporal delays) does not uniquely specify their spatial relationship at each temporal delay. Second, two Gabor functions (which were used to fit the spatial profiles of the vectors in Fig. 4) with a 90° phase difference are generally not orthogonal to each other. Even if the two vectors are orthogonal at each temporal delay (which is not imposed by the method), their phase difference still may not be 90°. Thus, the spatial phase relationship we have shown is not a trivial consequence of the method but is a reflection of the response property of complex cells. Another question is whether this method allows identification of visual features that are not orthogonal to each other. Generally, even if the features are not orthogonal, this method can still be used to identify linear combinations of the features. Subsequently, the relationship between the visual features and the neuronal response may be revealed by measuring the joint contrast–response function of the significant eigenvectors (see Fig. 6A). Finally, although PCA is a linear method for decomposing each stimulus into the sum of multiple eigenvectors, it does not require additive interaction between different eigenvectors in the response of the neuron. Even if the cell does not sum the responses to different features, this method can still identify either the individual features or linear combinations of them. The type of interaction between visual features can then be determined through analysis of the joint contrast–response function. These points can be demonstrated using simulated responses of model cells with the feature selectivity described above (data not shown). Finally, it is important to keep in mind that this method does not necessarily identify all of the features that affect the responses of the neuron, especially those that contribute weakly to the response.

Contrast–response function. For measuring the contrast–response functions, the contrast of the kth eigenvector in the stimulus, V_k, is defined as the dot product between the stimulus vector and the eigenvector, as follows:

$$\frac{1}{16} \sum_{t=1}^{16} \sum_{x=1}^{16} V_k(x,t) S(x,t),$$

where $-1 \le S(x,t) \le 1$ represents luminance at the tth temporal frame in the xth bar position of the stimulus pattern. Because the eigenvectors are normal:

$$\sum_{t=1}^{16} \sum_{x=1}^{16} |V_k(x,t)|^2 = 1,$$

the scaling factor 1/16 in the definition ensures that the contrast of each stimulus pattern is bound between -1 and 1. In the joint contrast–response function, the contrasts of both eigenvectors (Figs. 6A,B, 7A, contrast 1 and contrast 2) are defined the same as above.

Estimation of upper limit for correlation coefficient. To estimate the upper limit for the correlation coefficient between the predicted and measured contrast–response functions of relevant visual features (see Fig. 6), we simulated the functions measured from a finite number of repeats using a parametric bootstrap (Efron and Tibshirani, 1993). Briefly, for each stimulus that was repeated L times in the experiment with recorded responses $r_1, r_2, \ldots, r_L$, we simulated the response by drawing random samples $(r_1', r_2', \ldots, r_L')$ from a Gaussian distribution with the same mean and variance as the recorded responses $(r_1, r_2, \ldots, r_L)$ and computed the average of the simulated responses, as follows:

$$\frac{1}{L} \sum_{i=1}^{L} r_i'.$$

Repeating this step for all of the contrast levels resulted in a simulated contrast–response function with a noise level comparable with that measured experimentally. We then computed the mean $\pm 95\%$ confidence interval (obtained from 500 simulations) of the correlation coefficient between contrast–response functions obtained in different trials of the simulation. This was used as an estimate of the upper limit for the
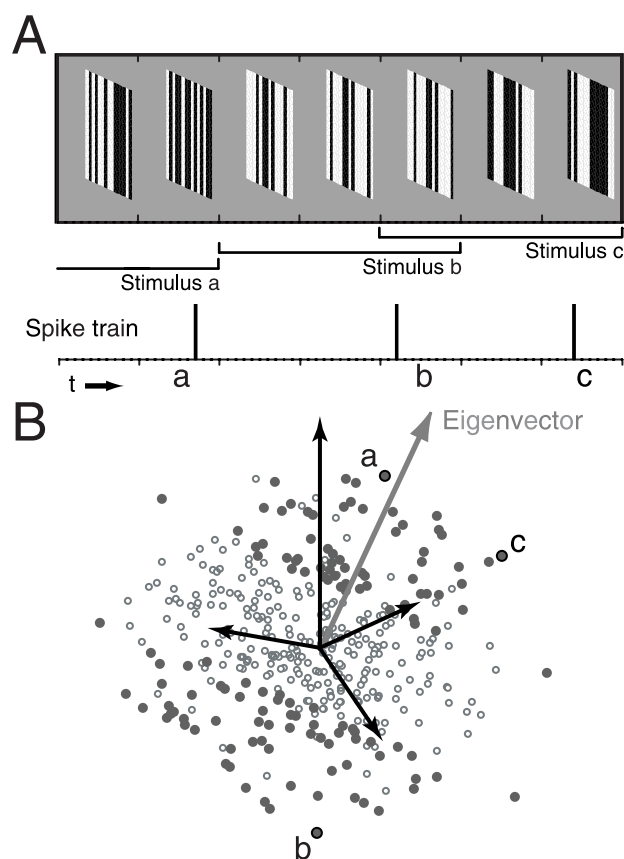
Figure 1. Illustration of a spike-triggered correlation analysis. *A,* Spatiotemporal random-bar stimuli (*top*) and the spike train response (*bottom*). Each *bracket* indicates a spatiotemporal stimulus pattern preceding a spike (for actual analysis, each pattern contained 16 rather than 3 frames). *B,* Schematic representation of the spike-triggered stimulus ensemble (*filled circles*) and the entire stimulus ensemble (*open* and *filled circles*) in a multidimensional parameter space. Each axis (*black arrows*) represents luminance at a particular bar position and time frame, and each *point* represents a stimulus pattern. Note that the actual stimulus ensemble is represented in a 256-dimensional space (16 frames by 16 bars). *a–c* indicate stimulus patterns shown in *A.* The *gray arrow* indicates an eigenvector of the spike-triggered ensemble with its eigenvalue (variance) greater than the eigenvalues of the entire ensemble.

correlation coefficient between the predicted and measured contrast–response functions set by noise in the measured responses.

RESULTS

Segregation between two types of visual features

Single-unit recordings were made from complex cells in the striate cortex of anesthetized adult cats. The stimuli consisted of 16 bars along the preferred orientation of the cell, with each bar varying randomly between light and dark at 60 Hz (Fig. 1*A*). To construct a basis set for each neuron that isolates the relevant visual features, we collected the spatiotemporal visual signals within a window of 268 msec (16 frames) before each spike and performed principal component analysis of this spike-triggered stimulus ensemble (Fig. 1*B, filled circle*) (see Materials and Methods). Unlike the spike-triggered average, which is the mean of the spike-triggered stimulus ensemble, the present method identifies a set of visual features (represented by eigenvectors of the spike-triggered correlation matrix) that account for different amounts of variance (the corresponding eigenvalues) in the ensemble. A visual feature with an outstanding variance (significantly larger or

smaller than the variance of the control ensemble) (Fig. 1*B, open circle*) is directly relevant to the spiking response of the neuron.

Figure 2*A* shows the 30 largest eigenvalues of the spike-triggered correlation matrix for a complex cell. Two eigenvalues (*filled circles*) conspicuously stood out from the rest (*open circles*), suggesting that the corresponding visual features (eigenvectors) are particularly relevant to the cell. The *dashed lines* indicate the confidence interval for eigenvalues of the control stimulus ensemble, sampled randomly from the random-bar stimuli (see Materials and Methods). The first two eigenvalues of the spike-triggered ensemble were well above the control, indicating significance of the corresponding eigenvectors. Figure 2*B* shows three eigenvectors, two corresponding to the significant eigenvalues (first and second) and one to a nonsignificant eigenvalue (*n*th). Although the spatiotemporal structure of the nonsignificant eigenvector appeared to be random, the significant eigenvectors had spatially separate on and off subregions evolving smoothly over time. To further confirm the distinction between these two types of eigenvectors, we compared both their eigenvalues and the correlation in their structures (which is a measure of nonrandomness) for a population of complex cells (*n* = 60). Figure 2*C* shows the distributions of the significant and nonsignificant eigenvalues; Figure 2*D* shows the distributions of the correlation of the eigenvectors (legend to Fig. 2). The two types of eigenvectors showed little overlap in both properties, indicating an unambiguous segregation between them.

For most (47 of 60) of the complex cells studied, we found two significant eigenvectors (Fig. 3), corresponding to the two largest eigenvalues. These two eigenvectors exhibited separate on and off spatial subregions (Fig. 2*B*), resembling the receptive fields of simple cells. The relationship between the two vectors was revealed by fitting their spatial profiles at the peak temporal delay (~40 msec preceding spike) with Gabor functions (Fig. 4*A*). In all cases, the Gabor fits for the two vectors exhibited similar spatial frequencies but a difference of ~90° in phase (Fig. 4*B*), reminiscent of the relationship between different subunits in the energy model for complex cells (Movshon et al., 1978; Pollen and Ronner, 1981; Adelson and Bergen, 1985; Heeger, 1991). As explained in detail in Materials and Methods, this phase relationship reflects the response property of complex cells and is not a trivial consequence of the orthogonality between eigenvectors, which is imposed by the method. In a few cases (3 of 60), we found only one significant eigenvector for each complex cell; these vectors also exhibited spatiotemporal profiles resembling simple-cell receptive fields. In the remaining cases, more than two eigenvalues reached significance. However, these additional eigenvectors (corresponding to third, fourth, . . . , largest eigenvalues) tended to exhibit much less spatiotemporal structure than the first two eigenvectors, and their eigenvalues were much smaller, suggesting less functional importance.

Responses of cortical neurons to individual visual features

The clear segregation between the significant and nonsignificant eigenvalues suggests that the corresponding eigenvectors contribute differently to the cortical responses. To test this idea directly, we measured the responses of each complex cell to individual vectors in both categories. Each vector (a 268 msec movie) was presented at a range of positive and negative contrasts (see Materials and Methods for the definition of contrast), and the peristimulus time histograms (PSTHs) of the cell were measured (Fig. 5*A*). Note that only the last bin (indicated by an *arrow*) of

Figure 2. Distinction between the significant and nonsignificant eigenvectors. *A,* The 30 largest eigenvalues of the spike-triggered correlation matrix of a complex cell. *Dashed line,* Control confidence interval ($p = 10^{-4}$) obtained by random sampling of the entire stimulus ensemble (100 repeats; see Materials and Methods). *Filled circles,* Eigenvalues that are significantly different from the control; *open circles,* nonsignificant eigenvalues. *B,* Two significant eigenvectors (first and second) and one nonsignificant eigenvector (*n*th; $n = 15$ in this case) whose eigenvalues are indicated by the *large circles* in *A. Arrow,* 40 msec (the delay at which the spatial profiles of the eigenvectors are shown in Fig. 4*A*). Calibration: 1°, 100 msec. *C,* Distributions of significant (*sig.; solid line*) and nonsignificant (*nonsig.; dashed line*) eigenvalues from 60 complex cells. Each eigenvalue was normalized by the mean eigenvalue of the cell. *D,* Distribution of temporal correlation in significant and nonsignificant eigenvectors. The autocorrelation function of each eigenvector was computed along the temporal axis; the eigenvector correlation

shown here was measured by the autocorrelation at the delay of 1 frame (16.7 msec) normalized by the autocorrelation at 0 delay, as follows:

$$(\sum_{x=1}^{16} \sum_{t=1}^{16} V(x,t-1)V(x,t)/ \sum_{x=1}^{16} \sum_{t=1}^{16} V(x,t)V(x,t) + 1)/2,$$

where $V(x,t)$ represents the luminance value of the eigenvector at pixel x and temporal delay t (see Materials and Methods). This value provides a measure of the nonrandomness of the eigenvector structure between 0 (completely random) and 1 (temporally uniform).
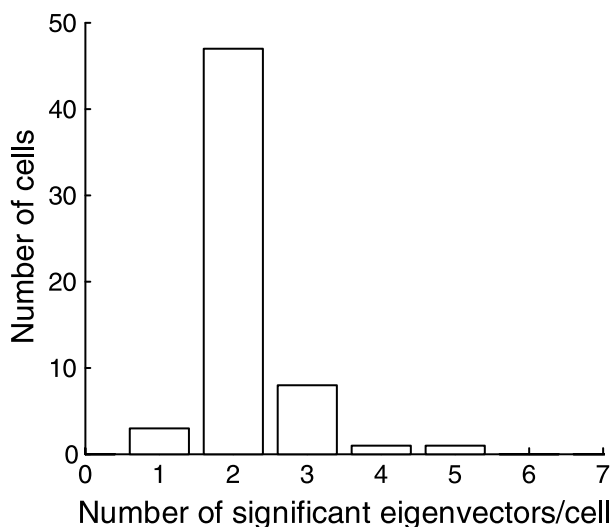

Figure 3. Distribution of the number of significant eigenvectors found for each cell, from a total of 60 complex cells.

Figure 4. Relationship between the two most significant eigenvectors. *A,* Spatial profiles (*solid lines*) of the two significant eigenvectors shown in Figure 2*B* at 40 msec from spiking (*arrows*) and their Gabor fits (*dashed lines*). The two Gabor fits had a phase difference of 90.4°. *Dotted line* represents mean luminance. *B,* Distribution of the spatial phase difference between the two significant eigenvectors of each complex cell.

each PSTH reflects the neuronal response to the complete spatiotemporal visual feature represented by the eigenvector; its amplitude was used to measure the contrast–response function. Figure 5*B* shows the contrast–response functions of a complex cell for two significant eigenvectors and one nonsignificant eigenvector. For each significant eigenvector, the response increased with the absolute value of the contrast at both positive and negative polarities, consistent with the known polarity invariance
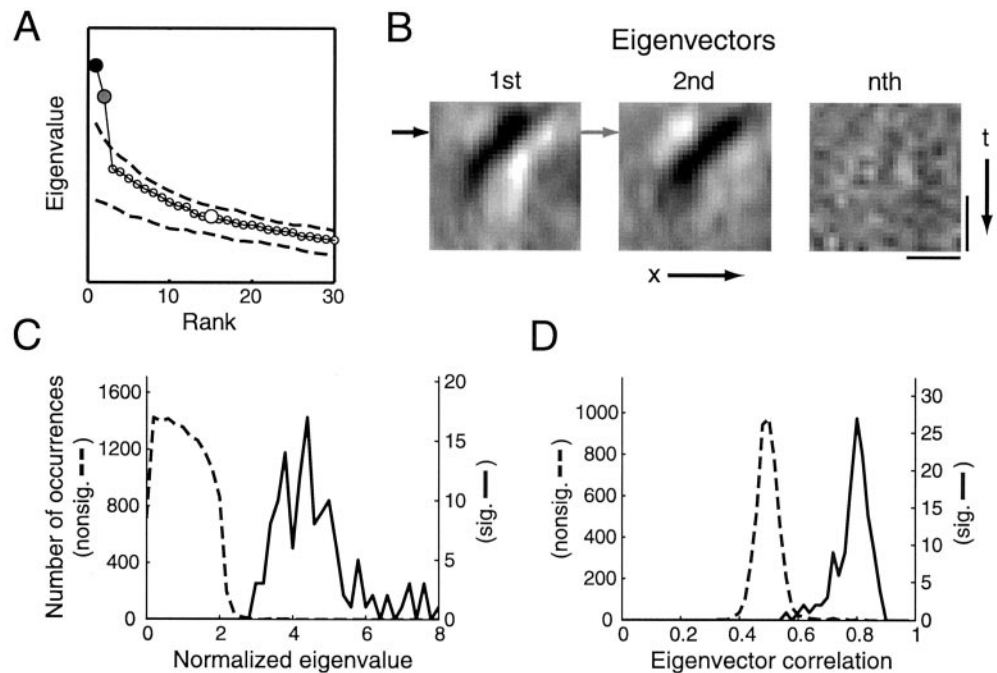
of complex cells (Hubel and Wiesel, 1962). The nonsignificant eigenvector, however, evoked no contrast-dependent response. We fitted the left and right sides of each contrast–response function separately with a power function, $y(x) = \beta|x|^\gamma$, where x and y represent the vector contrast and the neuronal response, respectively, and β and γ are free parameters. For the significant eigenvectors, the exponent γ was found to be 2.7 ± 0.1 (SEM; $n = 34$), similar to the exponent of contrast–response functions measured with drifting gratings (Albrecht and Geisler, 1991; Anzai et al., 1999). The ratio between the response at maximal vector
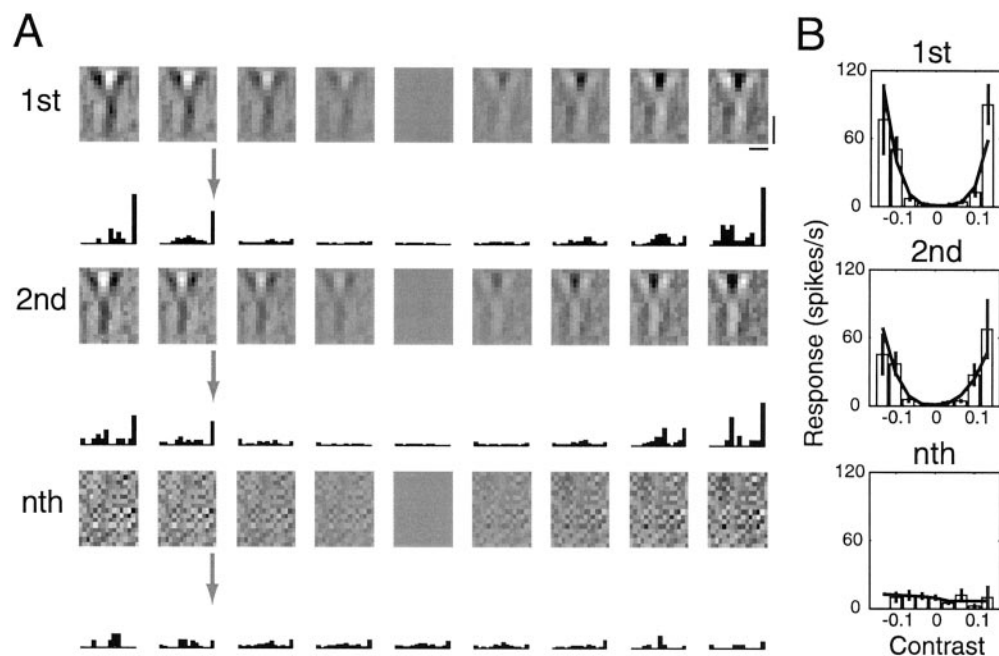
Figure 5. Responses of a complex cell (different from that shown in Fig. 2) to individual features. *A*, Responses to two significant eigenvectors (first and second) and one nonsignificant eigenvector (*n*th, randomly chosen for each cell; $n = 8$ in this case). The *top panel* of each row shows the eigenvector presented at a range of positive (*right*) and negative (*left*) contrasts. Calibration: 1°, 100 msec. The PSTH in response to each stimulus is plotted below, with the *arrow* indicating the last time bin of the PSTH. *B*, Contrast–response function (amplitude of the last bin in each PSTH) for each vector (*bars*) and the fit of each side with a power function (*line*). Error bars indicate SEM.

contrast and that at zero contrast was 96.6 ± 11.8. Thus, visual features represented by the significant eigenvectors can each drive the cortical neuron effectively in a contrast-dependent manner; they are referred to as relevant features. For the nonsignificant eigenvectors, the ratio between the responses at maximal and zero contrasts was 3.8 ± 0.9 ($n = 24$), much lower than that for the significant eigenvectors. Thus, visual features represented by the nonsignificant eigenvectors evoked little contrast-dependent response and were therefore termed null features.

Additive interaction between relevant visual features

Each spatiotemporal random-bar pattern can be decomposed into a combination of relevant and null features in the basis set. To understand cortical responses to arbitrary random-bar stimuli, it is necessary to characterize not only the contrast–response functions for individual features (Fig. 5) but also the interaction between features. First, we measured the responses of each complex cell to combinations of relevant features. For each neuron with two significant eigenvectors, we constructed a set of visual stimuli, each of which was a 268 msec movie consisting of a linear combination of the two significant eigenvectors. Figure 6*A* shows the responses of a complex cell at various combinations of the two vectors, which is referred to as the joint contrast–response function (see Materials and Methods). The response increased with the absolute value of contrast of either vector independently of their polarities, consistent with the contrast–response functions measured with individual vectors (Fig. 5*B*). Note that each combination of the two vectors also exhibited spatially separate on and off subregions (*small outer plots*), with the spatial phase shifting with the relative weights of the two vectors. The approximate circular symmetry of the joint contrast–response function indicates that the response is insensitive to the spatial phase of the stimuli, a well known property of cortical complex cells (Hubel and Wiesel, 1962; Movshon et al., 1978).

The approximate circular symmetry of the joint contrast–response function also suggests additive interaction between the two significant eigenvectors. To examine this idea quantitatively, we predicted the response to each combination of the two vectors

by summing the response to each vector at the corresponding contrast. As shown in Figure 6*B,* this prediction reproduced well the overall profile of the measured contrast–response function. Figure 6*C* shows the predicted responses (Fig. 6*B*) plotted against the measured responses (Fig. 6*A*) at corresponding contrasts; the correlation coefficient between them was found to be 0.84. To determine whether the difference between the predicted and the measured responses was attributable to systematic errors of the additive model or to the noise in the measured responses, we estimated the upper limit of the correlation coefficient set by noise in the responses. A contrast–response function measured in a single experiment was simulated with a Monte Carlo method, taking into consideration the variability of the measurement (see Materials and Methods); the contrast–response functions simulated in different trials were compared. As shown in Figure 6*D,* the correlation between the responses simulated in different trials (correlation coefficient, 0.87) was comparable with that between the predicted and measured responses, indicating that the additive model is consistent with the experimental results within the limit set by noise. Figure 6*E* summarizes the correlation coefficients between the measured and the predicted contrast–response functions for the 13 complex cells analyzed. In 12 of the cells, the correlation coefficient was not significantly different from that between simulated responses ($p > 0.05$), indicating that the model based on additive interaction provides an adequate description of the cortical responses to combinations of relevant visual features.

Divisive effect of the null features

Certain visual stimuli that do not evoke spiking responses on their own may nevertheless modulate cortical responses to other stimuli. Such nonlinear effects are well known for stimuli at nonpreferred orientations (Bonds, 1989) or nonclassical receptive fields (Allman et al., 1985; Walker et al., 2000) of cortical neurons. Here we tested whether the null features, which evoked little response when presented in isolation (Fig. 5), can modulate the cortical responses evoked by relevant features. The interaction between the relevant and null features was revealed by comparing

Figure 6. Interaction between relevant visual features. *A*, Joint contrast–response function of a complex cell (different from Figs. 2 and 5), in which the amplitude of the response (color-coded, with the scale shown on the *right* in units of spikes per second) is plotted against the contrasts of both relevant features. The *outer plots* depict the stimulus patterns corresponding to selected points (indicated by *arrows*) in the contrast–response function. Calibration: 1°, 100 msec. The range of contrast represented in the *center box* is −0.33 to 0.33; the contrast of each feature varied at a step size of 0.033. Responses were measured only at contrasts at which the luminance signals in the movie do not exceed the range of the monitor. *Gray* indicates contrast at which the response was not measured. *B*, Prediction of the contrast–response function based on additive interaction between the two vectors. The *left* and *bottom* histograms represent the contrast–response functions for the first and second eigenvectors, respectively, computed from the joint contrast–response function in *A* in the following manner: For eigenvector 1, the response at each contrast (contrast 1) was computed by averaging the measured re-

sponses across all values of contrast 2 (average across each row). Similarly, the contrast–response function for vector 2 (*bottom histogram*) was computed by averaging the responses in *A* across all values of contrast 1. Prediction of the response at each combination of contrast 1 and contrast 2 was then made by summing the value in each of the histograms at the corresponding contrast. Predictions were made only for contrasts at which the actual responses were measured in *A*. *C*, Measured response in *A* plotted against the additive prediction in *B* at corresponding contrasts. Each *circle* represents data at one pair of contrasts. *D*, Response from one simulation (*A*) plotted against that from another (*B*) (see Materials and Methods). *E*, Correlation coefficients between simulated responses plotted against the correlation coefficient between the predicted and measured contrast–response functions for 13 complex cells. *Vertical bars*, 95% confidence intervals of correlation coefficients between simulated responses, with the mean indicated by the *point* (the mean is not in the middle of the bar, because the distribution of the correlation coefficient is skewed).
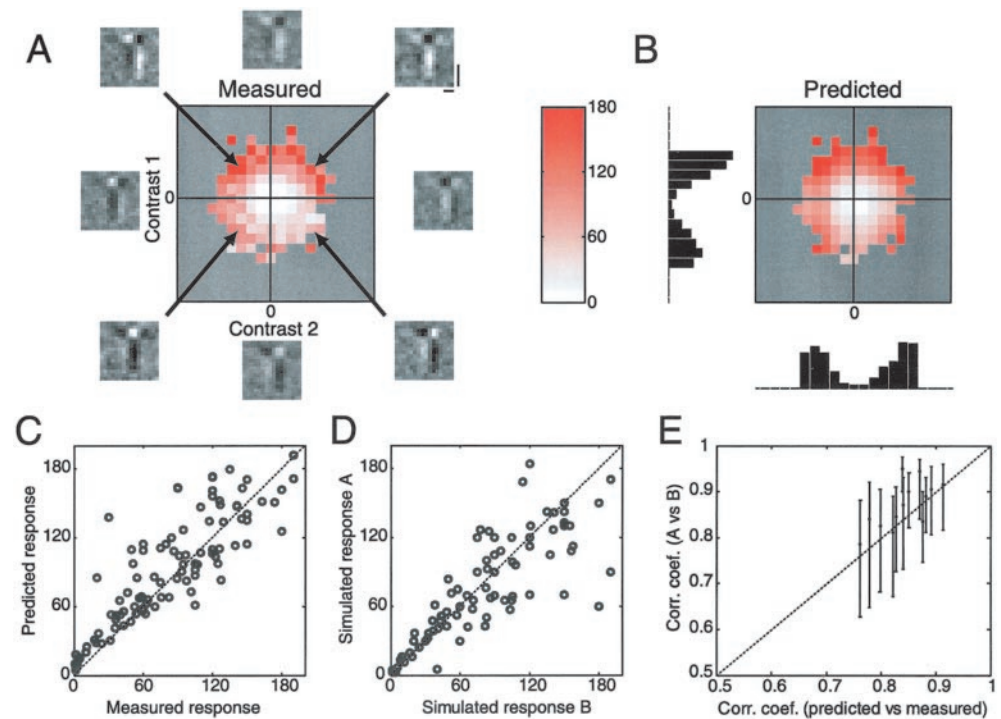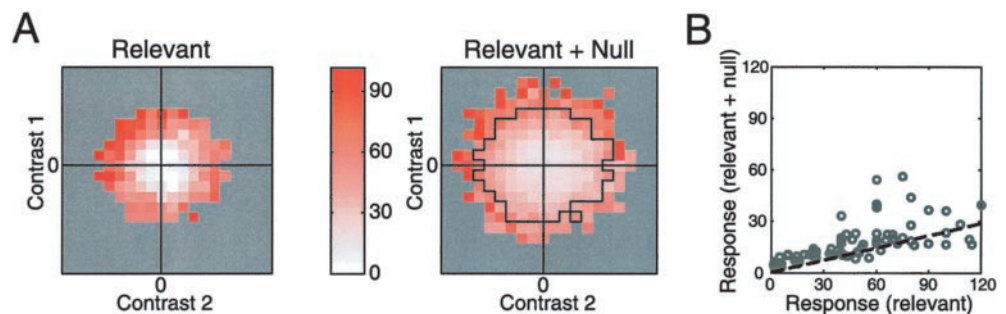
Figure 7. Suppressive effects of null features. *A*, Contrast–response functions of a complex cell for the two relevant features measured in the absence (*left*) or presence (*right*) of null features, with a luminance scale indicated in the *middle* (in spikes per second). The range of contrast represented in the *boxes* is −0.33 to 0.33. Contrasts of both features varied at a step size of 0.033. The *black line* in the *right plot* delineates the range of contrasts shown in the *left plot*. Higher contrasts for the relevant features were possible in the presence of null features, because the

superposition of certain null features can reduce the extreme luminance values in the short movie to levels within the monitor limit. *B*, Responses in the *left plot* versus the responses in the *right plot* in *A* at corresponding contrasts (in spikes per second). Each *circle* represents data at one pair of contrasts. The slope of the linear regression (*dashed line*) is 0.24.

the responses of each complex cell to relevant features alone (Fig. 6*A*) and to the random-bar stimuli (Fig. 1*A*) that contain both the relevant and the null features. Figure 7*A* shows the joint contrast–response function of a complex cell for the two relevant features, either in the absence (*left*) or presence (*right*) of null features. Although the two contrast–response functions exhibited similar shapes, the amplitude of the response to the random-bar stimuli was much lower (Fig. 7*B*), indicating a suppressive effect of the null features. Similar suppressive effects were observed for all of the cells examined.

The simplest models for this type of suppression are subtractive and divisive, and we evaluated both models in describing the effects of the null features. First, we fitted the contrast–response function for each relevant feature, either in the presence or in the

absence of null features, with power functions (Fig. 5*B*). The average scaling factor of the fit (β) was found to be 0.01 ± 0.002 ($n = 52$) in the presence of the null features and 0.03 ± 0.005 in the absence of them. The average exponents (γ) were 2.65 ± 0.13 ($n = 52$) and 2.95 ± 0.13 in the presence and absence of the null features, respectively. Although the null features reduced the scaling by a factor of ~3 ($p < 0.0005$; paired t test), they did not change the exponent systematically ($p > 0.10$). This is consistent with the observation that the null features changed the amplitude but not the shape of the contrast–response functions (Fig. 7*A*), suggesting a divisive effect. To compare directly the divisive and the subtractive models, we used both models to predict the joint contrast–response function measured in the presence of null features (Fig. 7*A*, *right plot*) from the function in the absence of
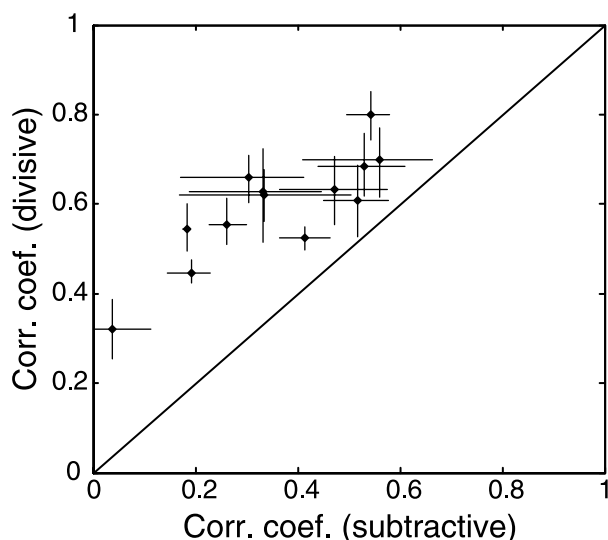
Figure 8. Comparison between the subtractive and divisive models for null features. Correlation coefficients between the measured responses and the prediction based on the divisive model were plotted against the correlation coefficients between the measured responses and the prediction based on the subtractive model. Each *point* represents the result of one cell. Error bar, 95% confidence interval, estimated using nonparametric bootstrap (see Materials and Methods for the Monte Carlo method used in the analysis).

null features (Fig. 7*A*, *left plot*). Each model contained a single free parameter (a scaling factor for the divisive model and a subtractive constant for the subtractive model) to ensure the fairness of the comparison. We found that for all of the cells analyzed ($n = 13$), the divisive model performed significantly better than the subtractive model ($p < 0.02$), as measured by the correlation coefficient between the predicted and measured responses (Fig. 8). Finally, we also fitted the predicted response based on the subtractive model with power functions. We found that the mean exponent of the fit was 5.60 ± 1.07 ($n = 52$), significantly larger than that for the measured responses ($p < 0.005$; paired t test). Together, these results support a divisive rather than a subtractive model for the suppressive effect of the null features.

DISCUSSION

In the present study, we have found that for each complex cell, visual inputs can be decomposed into two types of visual features, each having a distinct effect on the response of the cell. The two relevant features found for most complex cells resemble the receptive fields of simple cells, with a phase difference of ~90° in their spatial profiles; their contrast–response functions exhibited contrast polarity invariance and expansive nonlinearity reminiscent of a squaring function. Thus, the additive interaction between these relevant features corresponds closely to the energy model for complex cells (Movshon et al., 1978; Pollen and Ronner, 1981; Adelson and Bergen, 1985; Heeger, 1991). Although the energy model is well known, it has not been tested quantitatively with complex spatiotemporal stimuli in previous studies. The main difficulty in testing this model with complex stimuli comes from the fact that the parameters describing the underlying subunits of the energy model (simple-cell receptive fields) could not be determined easily for each cell. In the present study, relevant visual features were identified with spike-triggered correlation analysis, which allowed us to measure the contribution of

each feature to the cortical response directly and to demonstrate the additive interaction between them (Fig. 6). This result is also consistent with the finding that neural networks trained to predict the responses of complex cells to random-bar stimuli contained additive subunits resembling simple cells (Lau et al., 2002).

Divisive interactions have also been used to model the responses of both simple and complex cells (Heeger, 1992; Carandini et al., 1997); they can account for the suppressive effects of visual stimuli at nonpreferred orientations or nonclassical receptive fields of the cortical neurons. Such divisive suppression may reduce the redundancy in information carried by neighboring neurons and enhance the efficiency of coding for natural scenes (Schwartz and Simoncelli, 2001). Here, identification of a small number of relevant features for each cell allows us to specify the additive components in the visual inputs and to predict their contributions to the neuronal response. The number of null features that contribute to the suppression of cortical responses may be considerably larger. A similar spike-triggered analysis technique may be used to identify the null features that contribute maximally to the divisive suppression of the responses (Schwartz et al., 2001).

For sensory neurons with nonlinear stimulus–response relationships, it is often difficult to know *a priori* what visual stimuli are relevant for probing the response properties (Touryan and Dan, 2001). In the present study, we first isolated relevant features from null features for each cell using spike-triggered correlation analysis of the responses to a large ensemble of random spatiotemporal stimuli. This allowed us to construct new visual stimuli for each cell to measure the contribution of each type of features to the cortical response efficiently. Although this method has been used here to analyze the responses to random-bar stimuli, it is also applicable to studying cortical responses to more complex stimuli that vary in both dimensions of space (although with an increased number of parameters this analysis will require more data). Such a two-step approach may also prove to be useful for understanding the response properties of nonlinear neurons in other cortical areas and other sensory modalities.

REFERENCES

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2:284–299.
Albrecht DG, Geisler WS (1991) Motion selectivity and the contrast-response function of simple cells in the visual cortex. Vis Neurosci 7:531–546.
Allman J, Miezin F, McGuinness E (1985) Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. Annu Rev Neurosci 8:407–430.
Anzai A, Ohzawa I, Freeman RD (1999) Neural mechanisms for processing binocular information. I. Simple cells. J Neurophysiol 82:891–908.
Bell AJ, Sejnowski TJ (1997) The "independent components" of natural scenes are edge filters. Vision Res 37:3327–3338.
Bonds AB (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Vis Neurosci 2:41–55.
Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. Neuron 26:695–702.
Brodie SE, Knight BW, Ratliff F (1978) The response of the Limulus retina to moving stimuli: a prediction by Fourier synthesis. J Gen Physiol 72:129–166.
Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17:8621–8644.
Dan Y, Atick JJ, Reid RC (1996) Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. J Neurosci 16:3351–3362.
de Boer E, Kuyper P (1968) Triggered correlation. IEEE Trans Biomed Eng 15:169–179.
de Ruyter van Steveninck R, Bialek W (1988) Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and

information transfer in short spike sequences. Proc R Soc Lond B Biol Sci 234:379–414.

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. New York: Chapman and Hall.

Heeger DJ (1991) Nonlinear model of neural responses in cat visual cortex. In: Computational models of visual processing (Landy MS, Movshon JA, eds), pp 119–133. Cambridge, MA: MIT.

Heeger DJ (1992) Normalization of cell responses in cat striate cortex. Vis Neurosci 9:181–197.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol (Lond) 160:106–154.

Jones JP, Palmer LA (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. J Neurophysiol 58:1187–1211.

Lau B, Stanley GB, Dan Y (2002) Computational subunits of visual cortical neurons revealed by artificial neural networks. Proc Natl Acad Sci USA 99:8974–8979.

Marmeralis P, Marmeralis V (1978) Analysis of physiological systems: the white noise approach. New York: Plenum.

Movshon JA, Thompson ID, Tolhurst DJ (1978) Receptive field organization of complex cells in the cat's striate cortex. J Physiol (Lond) 283:79–99.

Olshausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381:607–609.

Pollen DA, Ronner SF (1981) Phase relationships between adjacent simple cells in the visual cortex. Science 212:1409–1411.

Reid RC, Victor JD, Shapley RM (1997) The use of m-sequences in the analysis of visual neurons: linear receptive field properties. Vis Neurosci 14:1015–1027.

Schwartz O, Simoncelli EP (2001) Natural signal statistics and sensory gain control. Nat Neurosci 4:819–825.

Schwartz O, Chichilnisky EJ, Simoncelli EP (2001) Characterizing neural gain control using spike-triggered covariance. In: Advances in neural information processing systems (Dietterich TG, Becker S, Ghahramani Z, eds). Cambridge, MA: MIT, in press.

Skottun BC, De Valois RL, Grosof DH, Movshon JA, Albrecht DG, Bonds AB (1991) Classifying simple and complex cells on the basis of response modulation. Vision Res 31:1079–1086.

Sutter EE (1987) A practical non-stochastic approach to nonlinear time-domain analysis. In: Advanced methods of physiological systems modeling (Marmarelis VZ, ed), pp 303–315. Los Angeles: University of Southern California.

Touryan J, Dan Y (2001) Analysis of sensory coding with complex stimuli. Curr Opin Neurobiol 11:443–448.

van Hateren JH, Ruderman DL (1998) Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. Proc R Soc Lond B Biol Sci 265:2315–2320.

Walker GA, Ohzawa I, Freeman RD (2000) Suppression outside the classical cortical receptive field. Vis Neurosci 17:369–379.

Wiener N (1958) Nonlinear problems in random theory. New York: Wiley.

Yamada WM, Lewis ER (1999) Predicting the temporal responses of non-phase-locking bullfrog auditory units to complex acoustic waveforms. Hear Res 130:155–170.