Visual Neuroscience (1997), **14**, 897–919. Printed in the USA.

Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification

WILSON S. GEISLER AND DUANE G. ALBRECHT

Department of Psychology and Center for Vision and Image Sciences, University of Texas, Austin

(RECEIVED July 13, 1996; ACCEPTED January 31, 1997)

Abstract

A descriptive function method was used to measure the detection, discrimination, and identification performance of a large population of single neurons recorded from within the primary visual cortex of the monkey and the cat, along six stimulus dimensions: contrast, spatial position, orientation, spatial frequency, temporal frequency, and direction of motion. First, the responses of single neurons were measured along each stimulus dimension, using analysis intervals comparable to a normal fixation interval (200 ms). Second, the measured responses of each neuron were fitted with simple descriptive functions, containing a few free parameters, for each stimulus dimension. These functions were found to account for approximately 90% of the variance in the measured response means and response standard deviations. (A detailed analysis of the relationship between the mean and the variance showed that the variance is proportional to the mean.) Third, the parameters of the best-fitting descriptive functions were utilized in conjunction with Bayesian (optimal) decision theory to determine the detection, discrimination, and identification performance for each neuron, along each stimulus dimension. For some of the cells in monkey, discrimination performance was comparable to behavioral performance; for most of the cells in cat, discrimination performance was better than behavioral performance. The behavioral contrast and spatial-frequency discrimination functions were similar in shape to the envelope of the most sensitive cells; they were also similar to the discrimination functions obtained by optimal pooling of the entire population of cells. The statistics which summarize the parameters of the descriptive functions were used to estimate the response of the visual cortex as a whole to a complex natural image. The analysis suggests that individual cortical neurons can reliably signal precise information about the location, size, and orientation of local image features.

Keywords: Primate visual cortex, Single neuron electrophysiology, Behavioral discrimination performance, Pattern recognition, Bayesian analysis

Introduction

The human visual system is able to detect, discriminate, and identify different patterns in space and time. The mechanisms underlying these abilities have been explored by studying the visual pathways of primates and cats. From this work, we know that neurons in the visual cortex respond to spatiotemporal contrasts, and that they are selective along a variety of stimulus dimensions, including contrast, spatial position, orientation, spatial frequency, temporal frequency, and direction of motion (e.g. Hubel & Wiesel, 1962, 1968; Campbell et al., 1968, 1969; for general reviews see: Robson, 1975; Orban, 1984; Shapley & Lennie, 1985; De Valois & De Valois, 1988; Palmer et al., 1991). To help understand the functional significance of the visual cortex neurons, their response selectivities have been evaluated within the context of behavioral psychophysics.

One strategy for comparing neural and behavioral performance is to measure the responses of single neurons using a stimulus protocol that is similar to a classical psychophysical protocol (Tolhurst et al., 1983; Bradley et al., 1985; Parker & Hawken, 1985; Barlow et al., 1987; Hawken & Parker, 1990; Geisler et al., 1991). An alternate strategy is to find simple functions which adequately describe the stimulus–response relationships of a given neuron and then use Bayesian decision theory to determine the performance (Geisler & Albrecht, 1995). This descriptive function method permits accurate determination of a neuron's performance with a smaller number of measurements, thereby making it feasible to conduct a broad survey of the performance characteristics of the visual cortex.

In this study, we measured cortical neuron responses in monkey and cat along a number of fundamental stimulus dimensions: contrast, spatial position, orientation, spatial frequency, temporal frequency, and direction of motion. We found that relatively simple functions provided an adequate description of the response means and the response variances. The parameters of these descriptive functions were used to determine the detection, discrimination, and identification performance along all of the stimulus dimensions, for a large population of neurons. The performance of this population of neurons was then compared with the behavioral performance of humans, monkeys, and cats. Finally, the statistics of the population were used to estimate the response of the cortex as a whole to complex natural images.

Reprint requests to: Wilson S. Geisler, Department of Psychology and Center for Vision and Image Sciences, Mezes Hall 330, University of Texas, Austin, TX 78712, USA.

Methods

Physiological preparation and stimulus display

The procedures for electrophysiological recording, stimulus display, and measurement of neural responses using systems analysis were similar to those described in detail elsewhere (Albrecht & Hamilton, 1982; Albrecht et al., 1984; Hamilton et al., 1989; Albrecht & Geisler, 1991), with the most recent changes described below. The protocols used in these experiments were approved by the University of Texas at Austin Institutional Animal Care and Use Committee and they conform to USDA regulations and NIH guidelines for the humane care and use of laboratory animals. The experiments were performed on macaque monkeys (*Macaca fascicularis*) and domestic cats. Anesthesia was induced using either ketamine hydrochloride (20 mg/kg/h) or thiopental sodium (20 mg/kg/h), and then maintained for the surgical procedures using thiopental sodium (20–30 mg/kg/h) for the cats, or sufentanil citrate (8–12 μg/kg/h) for the monkeys. During this time, a leg vein was cannulated for infusion; a stainless-steel pedestal was attached to hold the head in position without ear and eye bars; tissue was removed over the skull for a craniotomy; and the dura mater was thinned for insertion of a microelectrode. Following the surgical procedures, anesthesia was maintained for cats with continuous inhalation of nitrous oxide/oxygen (75%/25%) and continuous infusion of thiopental sodium (1 mg/kg/h), or with thiopental sodium alone (2–6 mg/kg/h); for monkeys, anesthesia was maintained with sufentanil citrate (4–6 μg/kg/h). Anesthesia and respiratory gases were monitored *via* an Ohmeda Respiratory Gas and Patient Monitor (Ohmeda RGM Model 5250, Critikon Dinamap Model 8300, Vet/Ox Pulse Oximeter Model 4402), which monitors the inspired and expired CO_2, O_2, N_2O, airway pressure, oxygen saturation in the blood, heart rate, and blood pressure.

The eyes were immobilized through continuous infusion of gallamine triethidiode (10 mg/kg/h) or pancuronium bromide (0.2 mg/kg/h). The animal was artificially respired through an endotracheal throat tube; expired CO_2 was maintained at 4.5% by adjusting the respirator rate and stroke volume. Accommodation was paralyzed and the natural pupil dilated by applying cyclopentolate hydrochloride. For cats, the nictitating membrane was retracted with phenylephrine hydrochloride. The eyes were covered with gas permeable hard contact lenses to prevent corneal drying. The animal was refracted using streak retinoscopy, corrective lenses were used to focus the stimuli on the retina, and a 3-mm artificial pupil was introduced. Core body temperature was continuously monitored and maintained *via* a thermostatically controlled hydraulic heating pad which surrounded the animal. Action potentials (collected in 0.1-ms time bins) were recorded from area 17 neurons using glass-coated platinum–iridium microelectrodes. The receptive fields for the neurons measured in this study were located within 5 deg of the visual axis.

The stimuli were luminance-modulated, drifting sine-wave gratings patterns for all of the stimulus dimensions except spatial position, where they were counterphase gratings. The grating stimuli were smoothly damped at the edges in space with abrupt onsets and offsets in time. They were presented on a Conrac studio monitor running at a noninterlaced frame rate of 100 Hz. The mean luminance was held constant at 27.4 cd/m^2. Both hardware and software methods were utilized to compensate for nonlinearities in the stimulus display. Responses were measured as a function of contrast, spatial position, orientation, spatial frequency, temporal frequency, and direction of motion, presented in a pseudorandom fashion. A single presentation at a fixed spatiotemporal contrast consisted of a block of ten or 12 contiguous temporal cycles. Each block was separated by a period of time equal to the block length; during these separations, the contrast was zero. Generally, four blocks were obtained for each stimulus condition, which resulted in 40 or 48 repeated temporal cycles (in some instances as many as 20 blocks were obtained). To assess the detection, discrimination, and identification performance within an analysis interval roughly comparable to a single fixation, the total number of spikes for each cycle was expressed in terms of spikes per 200 ms.

Analysis

Descriptive functions for the mean and variance

The measured means and standard deviations for each stimulus dimension were fitted using the descriptive functions shown in Table 1. In this table, we represent the relative response along each stimulus dimension by a selectivity function normalized to a peak of one. The mean of the response (r) is equal to the normalized function (r_x) times the maximum response (r_{max}) plus the base rate (r_0). The variance of the response (σ^2) is equal to the response mean times a constant (K). No special significance should be attached to these descriptive functions other than the fact that they provide a good fit to the measured responses with a minimal number of free parameters. These functions are simple and yet flexible enough to handle the wide variety of tuning function shapes.

Consider first the functions that describe the mean. For the dimension of contrast, the function is a saturating power function (the

Table 1. *Descriptive functions for the response means and variances*

Response mean	$r = r_{max} \cdot r_x(x) + r_0$
Response variance	$\sigma^2 = K \cdot r$
Contrast	$r_c(c) = \dfrac{c^n}{c^n + c_{50}^n}$
Spatial frequency	$r_u(u) = \begin{cases} \exp\left(-\ln 2 \dfrac{(u-u_c)^2}{h_u^2}\right) & u \geq u_c \\ \exp\left(-\ln 2 \dfrac{(u-u_c)^2}{l_u^2}\right) & u < u_c \end{cases}$
Temporal frequency	$r_w(w) = \begin{cases} \exp\left(-\ln 2 \dfrac{(w-w_c)^2}{h_w^2}\right) & w \geq w_c \\ \exp\left(-\ln 2 \dfrac{(w-w_c)^2}{l_w^2}\right) & w < w_c \end{cases}$
Spatial position	$r_p(p) = \alpha \lvert \sin[2\pi(p - p_0)] \rvert^n + 1 - \alpha$
Orientation	$r_\theta(\theta) = \exp\left(-\ln 2 \dfrac{(\theta-\theta_c)^2}{b_\theta^2}\right)$
Direction of motion	$r_d(d) = \begin{cases} 1 & \text{if } d = 1 \quad \text{(preferred)} \\ \rho & \text{if } d = -1 \quad \text{(nonpreferred)} \end{cases}$

c = contrast; p = position; u_c = center spatial frequency; w_c = center temporal frequency; θ_c = center orientation; p_0 = null position; n = response exponent; K = variance proportionality constant; u = spatial frequency; θ = orientation; h_u = spatial bandwidth (high); h_w = temporal bandwidth (high); b_θ = orientation bandwidth; α = relative null amplitude; r_{max} = response maximum; w = temporal frequency; d = direction of motion; l_u = spatial bandwidth (low); l_w = temporal bandwidth (low); ρ = direction selectivity ratio; c_{50} = half saturation contrast; r_0 = base response

Naka-Rushton function). For the dimensions of spatial frequency and temporal frequency, the function is a skewed Gabor function (a Gaussian whose standard deviation is allowed to differ on either side of the peak). For the dimension of spatial position, the function is a rectified sine wave raised to an exponent, plus a vertical offset. For the dimension of orientation, the function is a Gaussian. Finally, for direction of motion, the function is a constant representing the relative attenuation in the nonpreferred direction.

Consider next the function which describes the variance. This is a very simple relationship which states that the variance is proportional to the mean. In other words, the variance is predicted by one parameter—the proportionality constant K. Previous investigators have used a two-parameter function, where the variance is proportional to the mean raised to an exponent (e.g. Tolhurst et al., 1983). In Appendix B, we show that the simpler one-parameter function describes the measured responses just as well as the two-parameter function.

The descriptive functions for the means and standard deviations were fitted simultaneously using a maximum-likelihood procedure. For example, for the dimension of contrast, five parameters were estimated simultaneously: two for the normalized function, one for the maximum response, one for the base rate, and one for the variance. When more than one dimension was fitted simultaneously (Figs. 2 and 4), three of the parameters (r_{max}, r_0, K) were not allowed to vary across dimensions. The details of the maximum-likelihood fitting procedure are given in Appendix A.

Detection and discrimination performance

The descriptive functions fitted to each neuron's responses were used to determine what the neuron's detection and discrimination performance would be in a standard two-interval, two-alternative, forced-choice task. In this task, there are two possible stimuli. In a single trial, both stimuli are presented (in separate temporal intervals) in a random order. The goal is to indicate correctly which stimulus appeared in the first temporal interval. Threshold is defined as the difference in the two stimuli which gives 75% correct performance.

Thresholds for individual neurons were found using Bayesian decision theory (ideal-observer theory). We chose to use Bayesian decision theory because it shows the best performance possible given the neural responses, and hence provides a useful benchmark. To maximize the percentage of correct responses, the Bayesian observer uses (to close approximation) the following decision rule: if the response in the first interval is greater than the response in the second interval, then guess that the first interval contained the stimulus which produces the larger mean response; otherwise, guess that the second interval contained the stimulus which produces the larger mean response. The threshold achieved with this decision rule can be found using the formula:

$$d' = \frac{|\Delta Mean|}{\sqrt{Average\ Variance}} \tag{1}$$

The quantity d' is the absolute value of the difference in the means divided by the standard deviation. A d' of 1.0 corresponds to 75% correct in the two-interval forced-choice task. In other words, to discriminate two stimuli with an accuracy of 75%, the change in the mean response must equal one standard deviation. Substitution of a descriptive function for the response mean and variance into eqn. (1) results in the following formula for detection and discrimination performance:

$$d' = \frac{|r(x + \Delta x) - r(x)|}{\sqrt{\frac{K \cdot r(x + \Delta x) + K \cdot r(x)}{2}}} \tag{2}$$

where Δx is the change along the stimulus dimension. Detection and discrimination thresholds were found by setting d' to 1.0 and solving for Δx (i.e. incrementing or decrementing Δx from 0 until d' equaled 1.0).

Bayesian decision theory was also used to determine thresholds for the entire population of neurons as a whole, under the assumption that the responses of the neurons are statistically independent. The Bayesian decision rule (to close approximation) is to compute a weighted sum of the differences in the responses between the first and second stimulus intervals. If the weighted sum is positive one of the alternatives is picked, if it is negative the other alternative is picked. The d' for the entire population (using this decision rule) can be determined from the d's for the individual neurons in the same task. Specifically, the population d' is given by the following formula (Green & Swets, 1974):

$$d' = \sqrt{\epsilon \sum_{i=1}^{m} d_i'^2} \tag{3}$$

where d_i' is the value of d' for the ith neuron, m is the number of neurons in the population, and ϵ is the overall pooling efficiency.[1] The efficiency parameter quantifies the fraction of information utilized from each neuron; it varies from 0.0 (where none of the information is utilized) to 1.0 (where all of the information is utilized in an ideal fashion). Finally, the population threshold was obtained by setting d' to 1.0 and solving for Δx. For the purpose of comparing the shapes of behavioral and neural discrimination functions, the efficiency parameter, ϵ, was allowed to vary. To compare efficiencies across populations with different numbers of neurons, it is necessary to normalize for the number of neurons. To do this, we multiplied the best-fitting efficiency parameter, ϵ, by the number of cells, m, and then divided by 100 to obtain the efficiency per 100 cells, ϵ_{100}.

Identification performance

Identification involves making a decision about which stimulus was presented when the set of possible stimuli is large.[2] Bayesian decision theory was used to determine identification performance. The descriptive functions fitted to the measured responses were used to determine how accurately the stimuli could be identified based upon a single response. Specifically, we quantified identification performance by finding maximum-likelihood confidence regions, which define the most probable subset of stimuli given the observed response; the smaller the confidence region, the more accurate the identification performance. For example, a 95% confidence region contains, by definition, the correct stimulus 95% of

[1] Strictly speaking, the d' summation formula is derived assuming normal distributions for the responses, with the same variance for signal + noise and noise alone (the variance is allowed to be different for each neuron). However, if the average variance for signal + noise and noise alone is used in the computation of each d' [see eqn. (1)] then eqn. (3) is quite accurate for multiplicative noise (variance proportional to the mean) and for response distributions that are not normal (e.g. the Poisson distribution). We have shown this to be true in simulations where we compared the performance of the exact ideal observer with the performance predicted by eqn. (3).

[2] Detection and discrimination are special cases of identification where the number of stimuli is just two.

the time. If this region is small, the stimulus can be identified as belonging to that small region with 95% accuracy. This makes the confidence region a useful quantitative index of identification performance (Geisler & Albrecht, 1995).

To find the confidence region, we must determine the probability of the stimulus ($\mathbf{x}$) given the response (R), $p(\mathbf{x}|R)$. Bayes' formula expresses this probability in terms of the probability of the response given the stimulus, $p(R|\mathbf{x})$, and the prior probability of the stimulus, $p(\mathbf{x})$:

$$p(\mathbf{x}|R) = \frac{p(R|\mathbf{x})p(\mathbf{x})}{\int_{\mathbf{x}} p(R|\mathbf{x}')p(\mathbf{x}')\,d\mathbf{x}'} \tag{4}$$

where the integral in the denominator is over all possible stimuli $\mathbf{x}$. We assumed no knowledge of the stimulus prior to the response; that is, we assumed that $p(\mathbf{x})$ is constant, in which case, $p(\mathbf{x})$ cancels in the above equation. (Incorporating true stimulus probabilities, i.e. natural environmental constraints, should generally only improve identification performance.) The probability of the response given the stimulus was determined from the descriptive functions for the mean and variance under the assumption that the responses were approximately normally distributed:

$$p(R|\mathbf{x}) = \frac{1}{\sqrt{2\pi Kr(\mathbf{x})}} \exp\left\{-0.5\,\frac{[R - r(\mathbf{x})]^2}{Kr(\mathbf{x})}\right\} \tag{5}$$

This probability density function was then substituted into Bayes' formula to obtain the probability of the stimulus given the response.

$$p(\mathbf{x}|R) = \frac{\exp\left\{-0.5\,\frac{[R - r(\mathbf{x})]^2}{Kr(\mathbf{x})}\right\}}{\sqrt{r(\mathbf{x})}\int_{\mathbf{x}} \frac{1}{\sqrt{r(\mathbf{x}')}} \exp\left\{-0.5\,\frac{[R - r(\mathbf{x}')]^2}{Kr(\mathbf{x}')}\right\} d\mathbf{x}'} \tag{6}$$

This formula, when regarded as a function of the stimulus ($\mathbf{x}$), is referred to as the certainty function, or the *a posteriori* probability density function. Note that the shape of the certainty function depends upon the value of R.

Once the certainty function has been determined, the maximum-likelihood confidence interval is obtained by summing stimulus probabilities until the criterion confidence level, α, is reached. Formally, this process is equivalent to solving the following equation for β:

$$\alpha = \int_{\{\mathbf{x}:p(\mathbf{x}|R)>\beta\}} p(\mathbf{x}'|R)\,d\mathbf{x}' \tag{7}$$

where β is some level of stimulus probability. For example, a 95% confidence region is derived by starting with $\beta = 0$, and then increasing β in small steps until the sum (the integral in this equation) equals 0.95. We refer to this final value of β as β_{95}. The 95% confidence region, CR_{95}, is the set of stimuli whose probability density exceeds β_{95}; that is,

$$CR_{95} = \{\mathbf{x}: p(\mathbf{x}|R) > \beta_{95}\} \tag{8}$$

The smaller this set, the better the identification performance.

The above procedure is illustrated in Fig. 1. The solid curve represents a hypothetical certainty function: a probability distribution for the stimulus given a particular response, $p(\mathbf{x}|R)$. The shaded region has an area of 0.95; it defines the confidence interval indicated on the horizontal axis. As β is increased from zero, the sum of the probabilities which exceed β decreases until the sum equals 0.95 (the shaded area), which occurs when $\beta = \beta_{95}$.

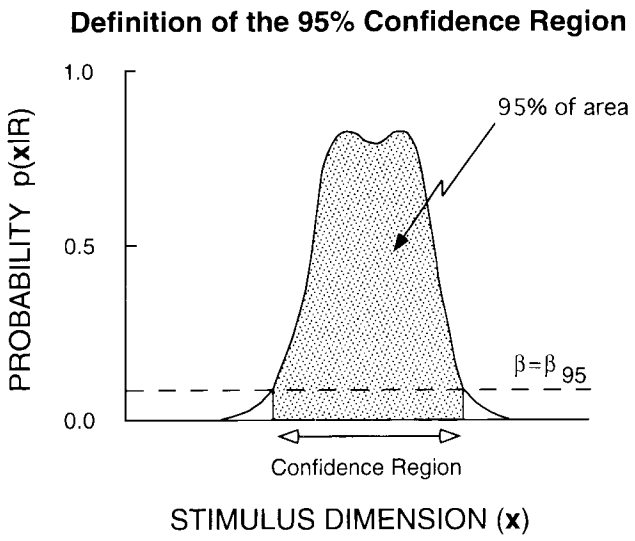

Fig. 1. Identification performance: The 95% confidence region. Definition of the maximum-likelihood 95% confidence region used to characterize the identification performance of single neurons. The smooth curve represents a certainty function (i.e. an *a posteriori* probability density function), which gives the probability of each possible stimulus ($\mathbf{x}$) given that a particular response (R) occurred during a single trial. The 95% confidence region (indicated by the two-headed arrow along the horizontal axis) is the region in stimulus space that corresponds to 95% of the area (the shaded area) under the certainty function. It is found by increasing the criterion β (the dashed line) until the sum of the probabilities greater than β equals 0.95. If a neuron produces a response of magnitude R, then subsequent brain mechanisms can be 95% certain that the stimulus came from this confidence region. The smaller this confidence region the better the identification performance of the neuron.

Advantages and limitations of the performance measures

The standard method for measuring neural performance is to measure counting distributions (pulse-number distributions) for each stimulus level and from these determine performance using Bayesian decision theory. The major advantage of the descriptive function method proposed here is that it is considerably more efficient: performance can be measured along a number of stimulus dimensions within a relatively short period of time. One reason the method is efficient is that the underlying pulse-number distributions are approximated by normal distributions, estimated from measurements of just two summary statistics: the response mean and variance. In Appendix C, we show that the estimates of performance are quite accurate even when the pulse-number distributions deviate substantially from normality. Another reason the method is efficient is that performance is determined from descriptive functions containing only a few parameters, and these parameters are estimated simultaneously from all of the data. In other words, the method makes use of the *a priori* knowledge that the means and variances of cortical cells can be described by specific descriptive functions with only a few parameters. This *a priori* knowledge reduces the amount of data required to estimate a cell's response characteristics.

We have compared the results of the descriptive function method with a more traditional signal detection method where entire pulse-number distributions are measured. We found that the estimates of performance from the descriptive function method were comparable to those from the more traditional method (see Appendix C).

One potential limitation of the descriptive function method used here is that it assumes that the parameters (e.g. r_{max}, K) are relatively stable through time; that is, the method does not incorporate

nonstationary factors. Although it might be possible to include such factors, they are unlikely to play a significant role. First, the descriptive functions accounted for approximately 90% of the variation in the measured responses (e.g. Figs. 3 and 4), leaving only a small amount of residual variation to be explained by other factors. Second, in a series of analyses designed to measure the degree of nonstationarity in striate neurons, we found little evidence of nonstationarity (see Appendix D in Geisler et al., 1991). Finally, in the present study we tested for systematic drifts in the parameters by fitting several data sets collected at different times with a common set of parameters as opposed to a unique set of parameters for each data set. The results of this analysis were consistent with what would be expected of a stable stochastic process.

It is important to note that this report only considers performance based upon rate information. We have developed methods that can incorporate rate and temporal pattern information (Geisler et al., 1991). Using these methods, we found that for visual cortex neurons, temporal pattern information appears to be important for discriminating stimuli that differ in some temporal parameter. The descriptive function method developed here can be extended to take into account temporal pattern information.

Results

Descriptive functions

Fig. 2 shows the responses of representative neurons along each of the stimulus dimensions measured in this study: contrast, spatial

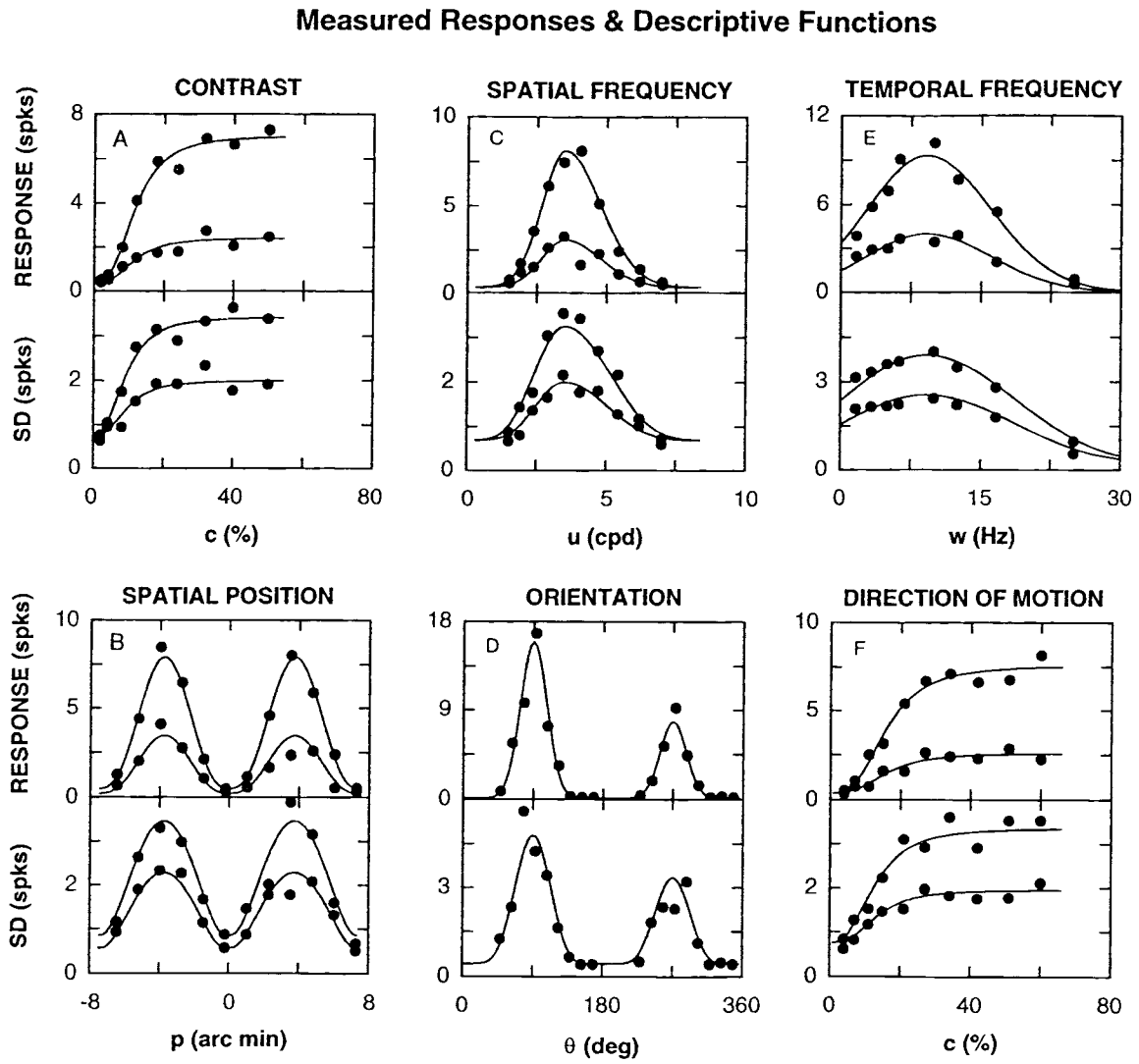

Fig. 2. Measured responses and descriptive functions. Responses of six monkey visual cortex neurons plotted as a function of six stimulus dimensions: contrast, spatial position, spatial frequency, orientation, temporal frequency, and direction of motion. For each dimension, the upper panel plots the mean response (the average total number of action potentials which occurred during a single 200-ms interval), and the lower panel plots the standard deviation associated with that response. Responses as a function of contrast (A) were measured at an optimal and a nonoptimal spatial frequency. Responses as a function of spatial position (B), spatial frequency (C), and temporal frequency (E) were measured at a saturating and a nonsaturating contrast. Responses as a function of orientation (D) were measured for motion in the preferred *versus* nonpreferred directions, at a saturating contrast. Responses as a function of contrast (F) were measured in the preferred *versus* nonpreferred directions for an optimal sine wave. The smooth curves through the data points are the fits of the descriptive functions given in Table 1. These functions provide a good description of the responses of most cortical cells for these stimulus dimensions.

position, spatial frequency, orientation, temporal frequency, and direction of motion. For each dimension, the upper panel plots the mean response, and the lower panel plots the standard deviation associated with that response. For example, in A, the responses are plotted as a function of contrast measured at two different spatial frequencies, and in B, the responses are plotted as a function of spatial position measured at two different contrasts. The smooth curves through the data points show the fits of the descriptive functions in Table 1 (see Appendix A for the maximum-likelihood fitting procedure). As can be seen, the descriptive functions provide a good fit to the measured responses. To quantify the goodness of fit, we used the following index:

$$G = 100 \times \left[1 - AVE\left(\frac{|observed - predicted|}{observed + predicted}\right)\right] \qquad (9)$$

A perfect fit would result in a value of 100%, and a very poor fit, a value of 0%.[3] The average values of the parameters of the descriptive functions for the entire sample of cells, along with the mean and median values of the goodness-of-fit index, are given in

[3] The ratio in this index is an error term that varies between 0 and 1 and is similar to a contrast index in that the difference is scaled by the mean. The average error is computed across all stimulus levels. Subtracting the average error from 1 and multiplying by 100 provides an index of the percentage of the variation accounted for. We chose this index because other indices are inadequate for describing the goodness of fit for complex cells in the spatial position experiment. Complex cells respond nearly the same to stimuli at different spatial positions. Conventional indices (e.g. 100 × [1 − residual variance/total variance]) indicate a very poor fit even when the data are well described as a constant increase in spike rate. For other stimulus dimensions, the index [eqn. (9)] is in good agreement with the conventional indices.

Table 2. *Population statistics for the descriptive functions*[a]

	Contrast						(247 cat 71 monkey)					
	n		c_{50}				K		G_r		G_σ	
Mean	3.0	2.5	25.0	43.0			1.7	1.7	89.0	89.0	90.0	90.0
Median	2.8	2.0	16.0	32.0			1.4	1.3	90.0	90.0	91.0	90.0
S.D.	1.4	1.4	25.0	31.0			1.0	1.2	6.2	6.2	4.4	4.9
	Spatial frequency[b]						(175 cat 96 monkey)					
	u_c		l_u		h_u		K		G_r		G_σ	
Mean	0.74	4.5	0.69	0.73	0.69	0.80	1.6	1.4	87.0	85.0	91.0	90.0
Median	0.62	4.2	0.67	0.66	0.65	0.78	1.4	1.2	88.0	87.0	91.0	90.0
S.D.	0.37	2.3	0.30	0.32	0.33	0.29	0.81	0.83	6.2	7.1	4.3	4.8
	Temporal frequency[b]						(119 cat 65 monkey)					
	w_c		l_w[c]		h_w		K		G_r		G_σ	
Mean	6.7	8.1	1.5	2.1	1.2	0.88	2.1	2.2	89.0	90.0	90.0	89.0
Median	6.1	7.2	1.2	1.7	1.3	0.71	1.9	1.7	90.0	91.0	91.0	90.0
S.D.	3.3	3.5	0.99	1.4	0.55	0.51	1.1	1.1	5.5	5.8	4.6	3.8
	Spatial position						(58 cat 27 monkey)					
	n						K		G_r		G_σ	
Mean	2.7	3.0					1.5	1.4	91.0	88.0	91.0	88.0
Median	2.3	2.4					1.3	1.2	92.0	90.0	92.0	89.0
S.D.	1.5	1.5					0.81	0.82	5.6	9.3	3.6	5.9
	Orientation						(21 cat 15 monkey)					
	b_θ		ρ				K		G_r		G_σ	
Mean	24.0	20.0	0.46	0.52			1.8	1.7	84.0	84.0	86.0	88.0
Median	21.0	15.0	0.45	0.51			1.7	1.4	81.0	82.0	85.0	88.0
S.D.	11.0	9.1	0.26	0.28			0.52	0.42	5.9	5.7	3.1	4.4
	Direction of motion						(47 cat 16 monkey)					
	ρ						K		G_r		G_σ	
Mean	0.62	0.46					1.6	1.7	86.0	94.0	88.0	94.0
Median	0.73	0.46					1.5	1.4	86.0	94.0	89.0	94.0
S.D.	0.30	0.26					0.75	0.58	6.5	1.9	4.6	1.3
	r_{max}		r_0									
Mean	8.0	8.2	0.36	0.52								
Median	5.7	5.7	0.09	0.17								
S.D.	6.6	7.0	0.66	0.83								

[a] For each parameter, cat statistics are in the left column and monkey the right.
[b] The bandwidths are expressed in octaves.
[c] For 41 cat cells and 32 monkey cells, the response did not fall to half of the maximum at the lowest frequency tested (at 0.25 Hz).

Table 2. (In this table, G_r is the goodness-of-fit index for the mean responses, and G_σ is the index for the standard deviations.) As can be seen, the values of the goodness-of-fit index are generally near 90%.

Some of the measurements and analyses reported in Fig. 2 and Table 2 are similar to those reported previously by us and by many other investigators (see references in the Introduction). In this study, the responses along all six stimulus dimensions were measured and analyzed using the same methods (e.g. the same mean luminance, trial duration, stimulus protocol), thereby facilitating comparison of the results.

One important aspect of the results is that the standard deviations are related to the means by a simple proportionality rule ($\sigma^2 = Kr$). To assess the adequacy of this rule, we estimated the value of K for each stimulus dimension individually using the maximum-likelihood procedure described in Appendix B, and quantified the goodness of fit to the standard deviation data using the index defined by eqn. (9). Fig. 3 shows the frequency distribution of the goodness-of-fit measure for each stimulus dimension. As can be seen, for all of the stimulus dimensions the values of the index are very high (greater than 90%). Thus, the simple proportionality rule captures most of the variability in the data.

A second important aspect of the results is that the value of K is essentially invariant for a given cell across stimulus dimensions. This remarkable fact is illustrated in Fig. 2, where the responses for each cell are shown for two stimulus dimensions. For example, in C, responses are plotted as a function of spatial frequency for two different contrasts, and in F, responses are plotted as a function of contrast for two different directions of motion. The solid curves through the standard deviations are the predictions for a single constant of proportionality. To assess the degree to which K is invariant across stimulus dimensions, we estimated the value of K for multiple stimulus dimensions simultaneously (see Appendix B); the number of dimensions varied from 2 to 6. Fig. 4 shows the frequency distribution of the goodness-of-fit index averaged across all of the stimulus dimensions for each cell in this sample. The median goodness of fit was greater than 90%. Thus, for each cell a single value of K describes the relationship between the mean and the variance across multiple stimulus dimensions.

Note that because the variance is proportional to the mean

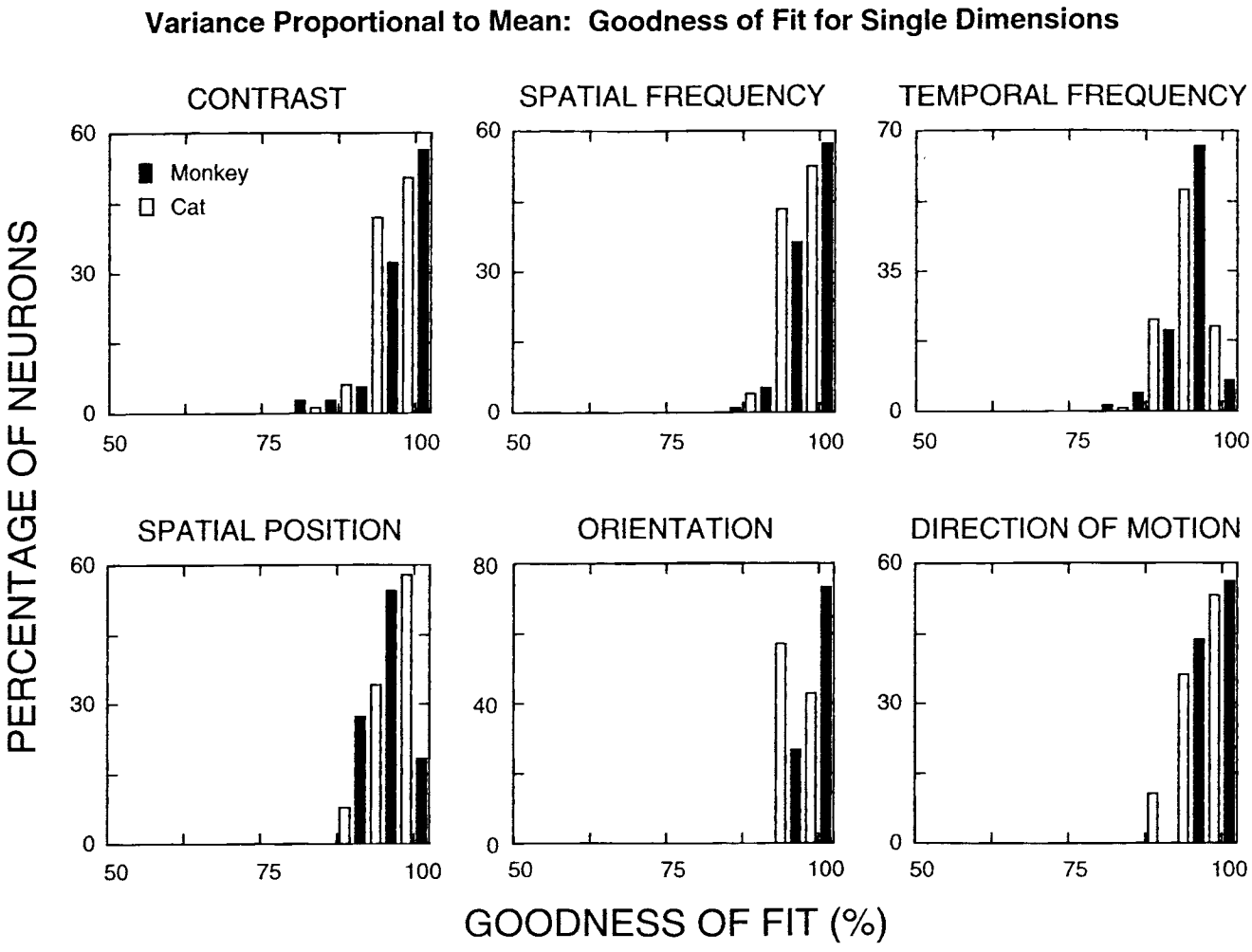

Fig. 3. Variance proportional to the mean: Single dimension. Histograms showing the goodness-of-fit index for the response standard deviations, G_σ, where the constant of proportionality, K, was estimated for each dimension individually. This goodness-of-fit index quantifies how well the one parameter proportionality rule, $\sigma^2 = Kr$, characterizes the measured standard deviations. The mean goodness-of-fit index for monkey was 93.7 (median = 94.4, S.D. = 3.2). The mean goodness-of-fit index for cat was 93.3 (median = 93.7, S.D. = 2.5). These high values indicate that the simple rule provides a good description of the variance across all of the stimulus dimensions examined. (See Appendix B for a more detailed analysis.)

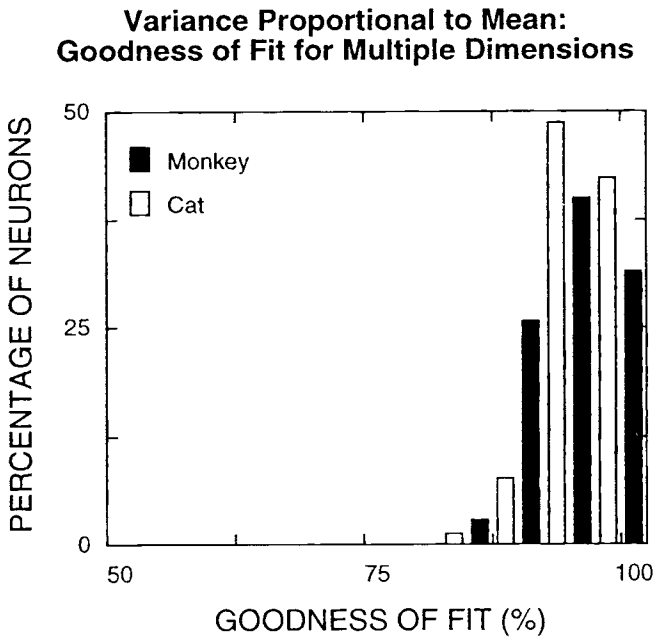

Fig. 4. Variance proportional to the mean: Multiple dimensions. Histogram showing the goodness-of-fit index for the standard deviation where the constant of proportionality, K, was fitted across more than one dimension simultaneously. This analysis was performed on a sample of 78 cat cells and 35 monkey cells. The mean goodness-of-fit index for monkey was 93.0 (median = 93.1, S.D. = 3.6). The mean goodness-of-fit index for cat was 94.4 (median = 94.6, S.D. = 3.0). These high values indicate that a single value of K holds across stimulus dimensions.

with a single value of K, then any two stimuli which produce equivalent mean responses will produce equivalent standard deviations. Consider, for example, the cell in Fig. 2D. If the stimulus is moving in the nonpreferred direction, at the optimal orientation, then the cell produces a mean response of approximately 8 spikes and a standard deviation of 3 spikes. If the stimulus is moving in the preferred direction, at a nonoptimal orientation approximately 20 deg from the peak, then the cell also produces a mean response of 8 spikes and a standard deviation of 3 spikes. Thus, the standard deviation is determined by the response independent of the stimulus.

One exception to the invariant-K rule is that for some cells K tends to increase at low temporal frequencies (below 2 Hz). This behavior can be seen in Fig. 2E; the observed standard deviations exceed the predicted at low temporal frequencies. For the majority of cells, K remains invariant as a function of temporal frequency. Further analysis of this effect is described in Appendix B.

In summary, for all of these fundamental stimulus dimensions, the relationship between the mean of the response and the variance of the response for any given cell is adequately described by a simple constant-ratio rule: the variance is proportional to the mean with a single proportionality constant that is independent of the stimulus dimension.

Detection and discrimination performance

As shown in the previous section, the response means and standard deviations of each neuron in the population were accurately fit by the descriptive functions in Table 1, thus making it possible to determine the detection and discrimination performance of each neuron.

Fig. 5 shows the discrimination functions for the neurons in Fig. 2. For example, in A, contrast threshold is plotted as a function of base contrast (i.e. background or reference contrast) for the optimal stimulus. This contrast-discrimination function is typical of those reported in the literature (for a review see Geisler et al., 1991). As base contrast increases, threshold initially decreases, reaches a minimum, and then rapidly increases. Thus, the detection threshold (the threshold at a base contrast of 0.0) is larger than the best discrimination threshold. Similar trends are seen for the cell in F, where contrast-discrimination functions are plotted for the preferred and nonpreferred directions of motion. In B–E, the discrimination functions are shown for the highest contrast measured.

Note that for each stimulus dimension good discrimination only occurs over a limited region. The smallest thresholds occur away from the peak, where the slope is steep and the response is still large (c.f. Fig. 2). For all of the stimulus dimensions except contrast, there are multiple regions of high sensitivity. This occurs because the response functions along these dimensions have multiple regions where the slope is steep.

It would, of course, be possible to improve discrimination performance by making the analysis intervals longer than 200 ms. However, intervals longer than 200 ms would not be comparable to fixation times during normal saccadic inspection. Further, even under steady fixation, longer intervals would exceed psychophysical estimates of the temporal integration limit of the visual system (Watson, 1986), and would thus overestimate the contribution of individual neurons to behavioral performance.

Fig. 6 and Table 3 summarize the detection and discrimination performance of the entire population of cells. Fig. 6 shows the frequency histograms of the best discrimination thresholds for each

Table 3. *Population statistics for cortical performance*

	Discrimination		Identification	
	Cat	Monkey	Cat	Monkey
Contrast	Δc (%)	Δc (%)	c (%)	c (%)
Mean	11.0	18.0	80.0	72.0
Median	5.2	12.0	83.0	74.0
S.D.	17.0	18.0	11.0	15.0
Spatial frequency	Δu (%)	Δu (%)	u (octaves)	u (octaves)
Mean	13.0	15.0	1.3	1.4
Median	12.0	14.0	1.2	1.4
S.D.	5.8	5.8	0.48	0.40
Temporal frequency	Δw (%)	Δw (%)	w (octaves)	w (octaves)
Mean	28.0	23.0	2.6	2.7
Median	25.0	19.0	2.6	2.8
S.D.	15.0	14.0	1.0	0.69
Spatial position	Δp (deg)	Δp (deg)	p (deg)	p (deg)
Mean	30.0	25.0	123.0	121.0
Median	5.2	24.0	130.0	116.0
S.D.	17.0	11.0	35.5	38.9
Orientation	$\Delta\theta$ (deg)	$\Delta\theta$ (deg)	θ (deg)	θ (deg)
Mean	14.0	11.0	54.0	45.0
Median	12.0	10.0	43.0	39.0
S.D.	9.3	6.6	35.0	24.0
Direction of Motion	Δc (%)	Δc (%)	r	r
Mean	14.0	12.0	0.82	0.75
Median	12.0	9.5	0.96	0.77
S.D.	12.0	4.6	0.20	0.19

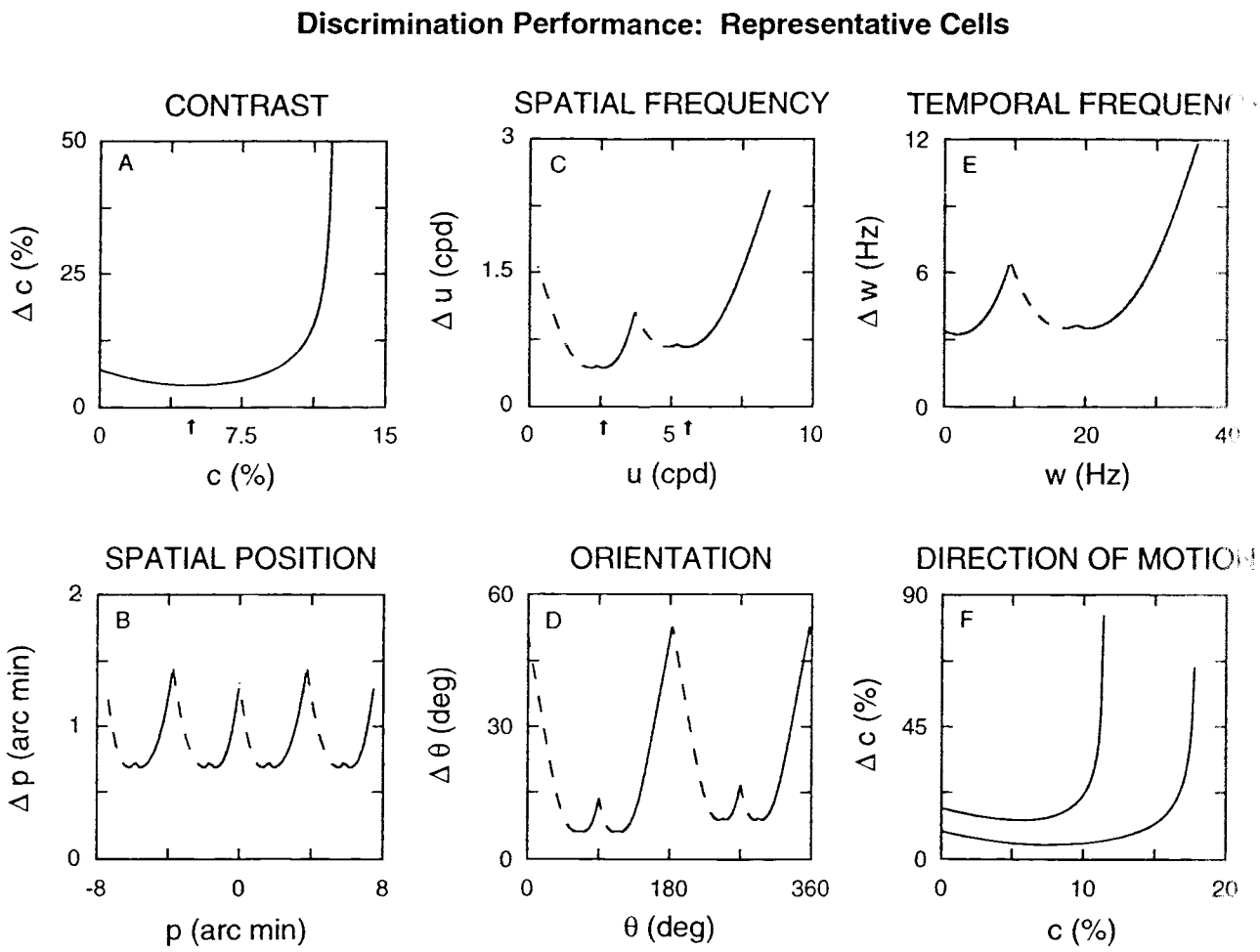

Fig. 5. Discrimination performance: Representative cells. Discrimination performance along each of the six stimulus dimensions for the neurons shown in Fig. 2. The descriptive functions in Fig. 2 were utilized to determine discrimination thresholds along each stimulus dimension. For spatial frequency, temporal frequency, spatial position, and orientation, the solid curves indicate those values along the horizontal axis where the decrement thresholds were smaller and the dashed curves indicate where the increment thresholds were smaller. The arrows along the horizontal axis for the dimensions of contrast and spatial frequency indicate where the best discrimination performance occurred.

of the stimulus dimensions.[4] Note that for position, the threshold is expressed in degrees of spatial phase because position discrimination is strongly dependent on the peak spatial frequency of the cell.

Identification performance

Identification performance for each neuron was quantified by measuring how accurately one could know the stimulus based upon the neuron's response. As described in the Methods section, the descriptive functions and Bayes' formula were used to determine the probability of each possible stimulus given a particular response; this is the certainty function. The curves in Fig. 7 show the certainty functions for the cells in Fig. 2, given that the cells produced a maximum response. For example, the certainty function for orientation is shown in D. As can be seen, the most probable stimulus (given a maximum response) is a grating moving in the preferred direction, restricted to a narrow band of orientations. The arrows on the horizontal axis indicate the region that sums to 95% of the area under the certainty function; this is the maximum-likelihood 95% confidence region. (The arrows on the vertical axis indicate β_{95}, the smallest probability included in the 95% confidence region; see Fig. 1.) Thus, when a maximum response is produced, a subsequent brain mechanism could know the orientation of the stimulus to within ± 19 deg, and the direction of motion, with 95% confidence. Similarly, the other certainty functions in Fig. 7 show that when a maximum response is produced one could know the spatial frequency to within ± 1.25 cycles/deg (cpd), the position to within ± 1.7 min of arc,[5] the contrast to within ± 42%, the

[4] Because complex cells produce nearly equivalent responses at all spatial positions, they were not included in the analysis of discrimination performance along the dimension of spatial position.

[5] The two separate confidence regions for spatial position reflect ambiguity in the spatial phase of the stimulus when the temporal phase of the response is ignored. For the simple cells analyzed here, the temporal phase shifts typically corresponded to intervals of 100 ms or more. Any subsequent brain mechanism which can detect these large phase changes would know which confidence region applies; hence, we report only the confidence interval for one of them. This convention was adopted for all of the cells in the analysis of spatial position.

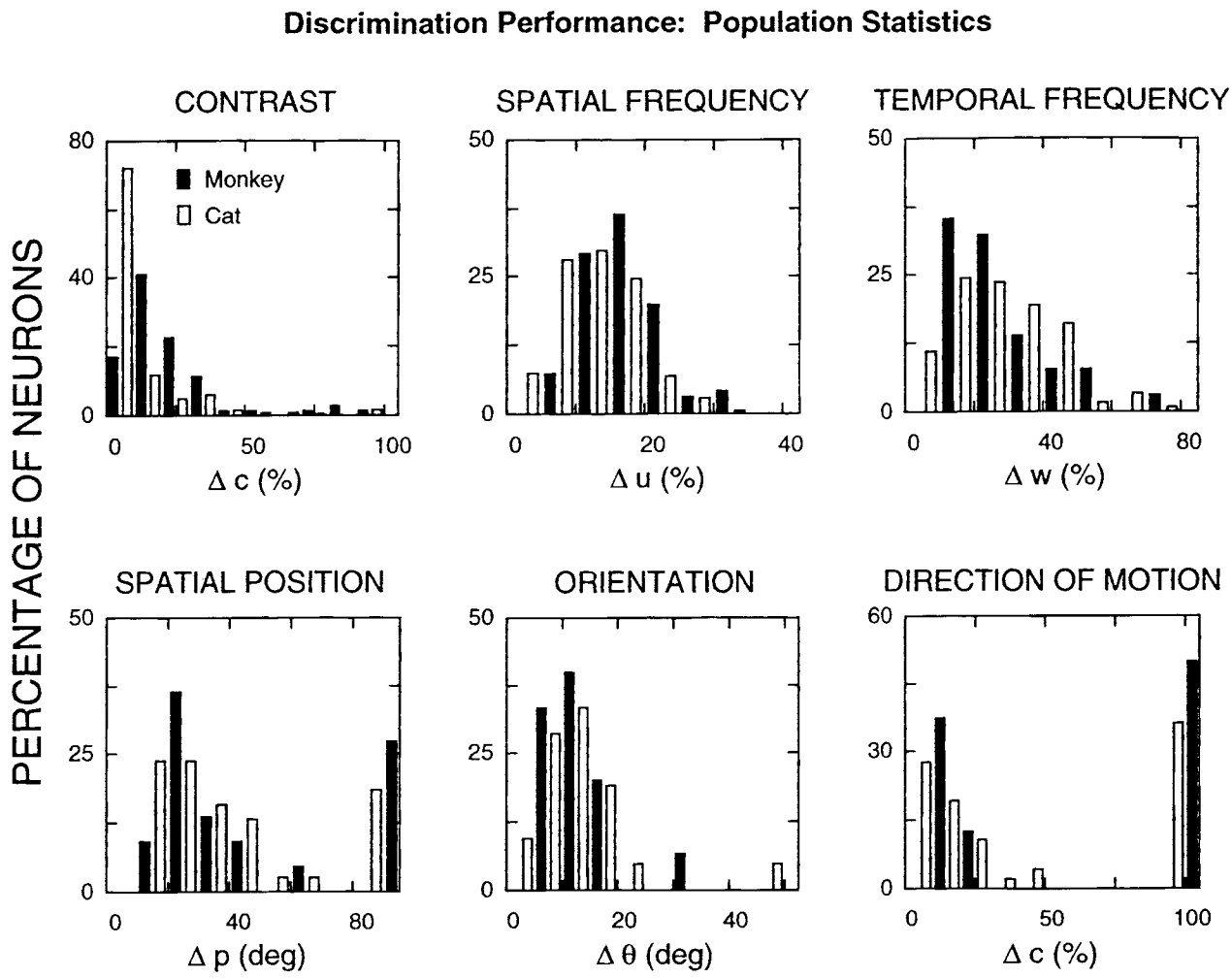

Fig. 6. Discrimination performance: Population statistics. Histograms showing the best discrimination performance for all of the cells, across all of the dimensions. For contrast, spatial frequency, and temporal frequency, the threshold is expressed as a percentage of the base or standard. For position the units are degrees of spatial phase. The means, medians, and standard deviations are given in Table 3. For direction of motion, the minimum contrast which evoked a reliably different response between the two directions was determined; the bins on the far right indicate cells which could not discriminate direction at any contrast. For spatial position, the bins on the far right indicate cells which could not reliably discriminate position/phase.

temporal frequency to within ± 7.8 Hz, and the direction of motion with greater than 95% confidence.

Fig. 8 and Table 3 show the 95% confidence regions along each of the stimulus dimensions for the entire sample of neurons. (Again, position is expressed in degrees of spatial phase because the confidence region is dependent on the peak spatial frequency of the cell.)

The width of the certainty functions and the size of the 95% confidence regions generally increase as the response decreases. Thus, for responses above the average maximum, the confidence regions would be smaller than those shown in Figs. 7 and 8, and for responses below the average maximum the confidence regions would be larger. However, the size of the confidence regions remains very similar for responses within approximately ±20% of the average maximum (Geisler & Albrecht, 1995).

Discussion

In an effort to understand the neurophysiological mechanisms of the visual system up to the level of the primary visual cortex, and how these mechanisms might contribute to visual performance, we measured the detection, discrimination, and identification performance of a large population of cortical neurons. These measurements were made along a number of fundamental stimulus dimensions, for analysis intervals comparable to the normal fixation duration. The method of measuring performance was to fit the response means and standard deviations with descriptive functions and then use these descriptive functions to determine performance. This method proved to be robust and efficient because the descriptive functions were able to account for a large percentage of the variation in the data with a relatively small number of parameters.

Response functions and performance

It is worthwhile to consider the relationship between the measured response functions and discrimination performance along each stimulus dimension (c.f. Figs. 2 and 5). Discrimination performance is good when the change in the mean response is large compared to the standard deviation. The change in the mean response is largest where the slope of the response function is greatest. On the other

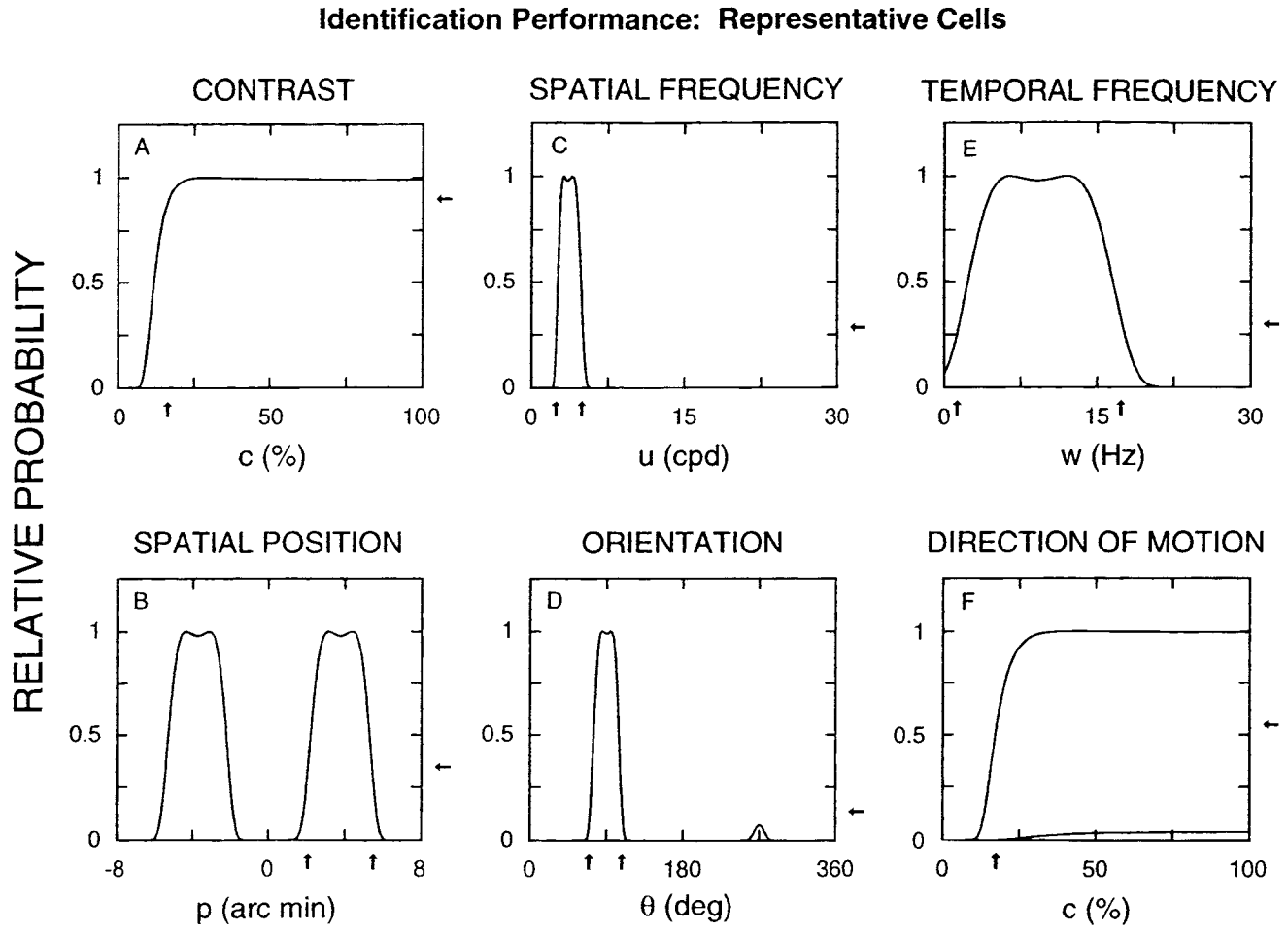

Fig. 7. Identification performance: Representative cells. Identification performance along each of the six stimulus dimensions for the neurons shown in Fig. 2. The descriptive functions were utilized to determine certainty functions: the probability of each stimulus, given that a maximum response has occurred. The certainty functions were then utilized to determine 95% confidence regions. The arrows along the horizontal axes indicate the 95% confidence region, and the arrows along the right vertical axes indicate the β_{95} for each dimension [see Fig. 1 and eqns. (7) & (8)]. If a maximum response occurs, subsequent brain mechanisms can be 95% certain that the stimulus is located between the arrows on the horizontal axis.

hand, the smallest standard deviations occur when the response is small, because the variance is proportional to the mean. Thus, the best discrimination occurs slightly below the steepest point of the response function. Above and below this point, performance decreases until either the bounds of the stimulus range are reached or the response function approaches a constant value and discrimination is no longer possible. For example, for the dimension of contrast (Figs. 2A and 5A) the slope of the response function is greatest at 9%, the best discrimination occurs at 5%, and discrimination becomes impossible beyond 12%.

To understand the relationship between the measured responses and identification performance, compare the means and standard deviations shown in Fig. 2 with the certainty functions shown in Fig. 7. Recall that the certainty functions in Fig. 7 show the probability of each particular stimulus given that a maximum response has occurred. As can be seen, the certainty functions are similar in shape to the response functions except that they have steeper falloffs and they are generally narrower. To understand why, consider a cell with a maximum response of 10 spikes and a variance proportionality constant of 1.5. A nonoptimal stimulus which produces a mean response of 2.5 spikes (25% of the maximum response) will produce a standard deviation of 1.9 spikes. Therefore, the maximum response of 10 spikes is four standard deviations above the mean response of 2.5 spikes, making it highly unlikely that this nonoptimal stimulus could ever produce the maximum response (the relative probability is less than 0.0003).

As a rule of thumb, if a stimulus produces less than half the maximum response then it falls outside of the 95% confidence region. In other words, the 95% confidence region corresponds approximately to the half-height bandwidth (c.f. compare Fig. 8 and Table 2). Given this rule, it is interesting to note that the polar plots in De Valois et al. (1982), which show the half-height bandwidths in spatial frequency and orientation, are approximately equivalent to the 95% confidence regions for these two dimensions, when a response near the maximum occurs.

Contrast and performance

In the primary visual cortex, the response functions along the various stimulus dimensions are relatively invariant in shape when measured at different contrasts, even at contrasts which produce response saturation (see for example, Figs. 2B, 2C, and 2E). This is because

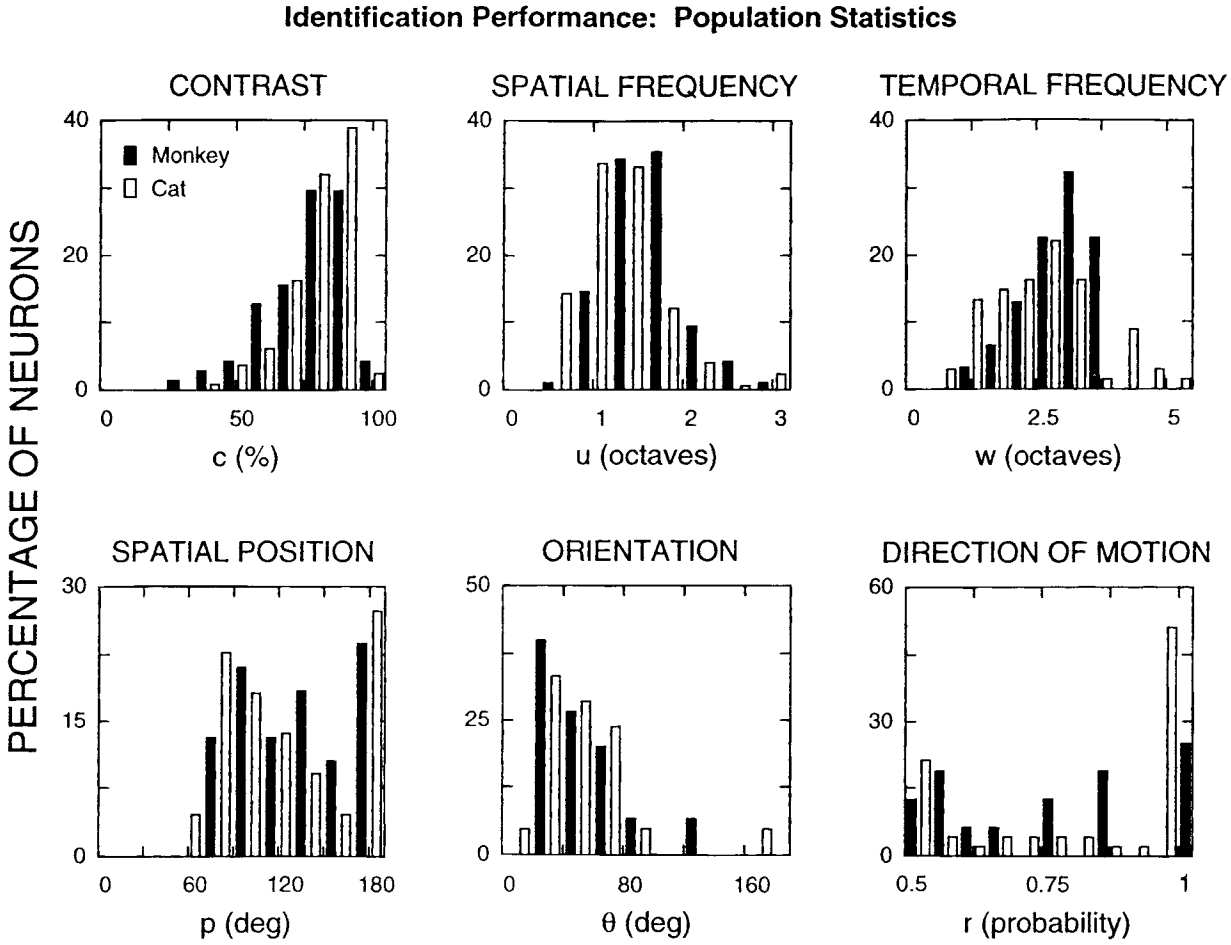

Fig. 8. Identification performance: Population statistics. Histograms showing the identification performance for all of the cells, across all of the dimensions. The means, medians, and standard deviations are given in Table 3. For all of the dimensions except direction of motion, the histograms represent the 95% confidence regions. For direction of motion, the histogram represents the probability that the stimulus was moving in the preferred direction. (For position, the units are degrees of spatial phase.)

response saturation in the cortex is not determined by the overall magnitude of the response but rather by the overall magnitude of the contrast (Albrecht & Hamilton, 1982; Sclar & Freeman, 1982; see Geisler & Albrecht, 1995, for a general review).

This fact implies that response saturation will not have a detrimental effect on discrimination for any dimension except contrast. This point is illustrated in Fig. 9, where contrast discrimination and spatial-frequency discrimination are plotted as a function of contrast for the same cell. When response saturation occurs (at approximately 15% contrast for this cell) contrast threshold becomes unobtainable whereas frequency threshold reaches its best value, and maintains that value at higher contrasts. Thus, the lower the contrast at which response saturation occurs, the better for frequency discrimination. Similar plots were obtained when contrast discrimination was compared with discrimination along the other stimulus dimensions.[6]

Similarly, response saturation does not have a detrimental effect on identification performance (except for contrast); in fact, cortical neurons provide their best identification performance when they are in the saturated region. As stimulus contrast increases and the response increases, the 95% confidence region decreases (Geisler & Albrecht, 1995). Hence, the smallest confidence regions will occur when stimulus contrasts are sufficient to produce maximum saturated responses. Further, because many cortical cells saturate at low contrasts (some as low as 5%), there will be many cortical cells providing good identification performance even for low-contrast image features.

Finally, we note that response saturation of the type observed in cortical neuron responses (i.e. contrast-determined response saturation) is not necessary for good discrimination performance; for example, good discrimination performance is possible with a linear mechanism as long as a small change in the stimulus produces a big change in the response. On the other

[6] If the response functions are invariant with contrast and the variance is proportional to the mean, then for any stimulus dimension, the signal-to-noise ratio (d') as a function of contrast will be proportional to the square root of the contrast response function. To see this, note that if the response function for a given dimension x is invariant with contrast then the joint two-dimensional response function can be described as the product of the response functions for x and for contrast: $r(x,c) = r_{max} r_x(x) r_c(c)$. The square-root relation follows by substituting $r(x + \Delta x, c)$ for $r(x + \Delta x)$, and $r(x,c)$ for $r(x)$, in eqn. (2). (The square-root relation would be expected to break down for small responses near the base rate r_0.)

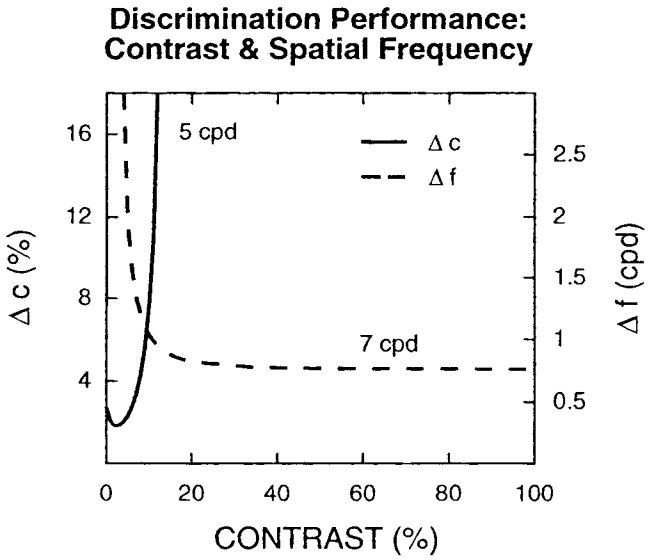

Fig. 9. Discrimination performance: Spatial frequency and contrast. Contrast discrimination and spatial-frequency discrimination plotted as a function of contrast for a neuron recorded from monkey visual cortex. The vertical axis on the left plots Δc (% contrast); the vertical axis on the right plots Δf (cycles/deg); the horizontal axis plots contrast. As can be seen, contrast discrimination is best from 0 to 10% contrast; whereas, frequency discrimination is best from 10% to 100% contrast, even though the cell saturated at 15% contrast.

hand, this type of saturation is essential for good identification performance because it insures that only those stimuli near the peak of the tuning function can produce big responses; good identification performance is not possible with a linear mechanism because large responses can be produced by nonoptimal stimuli if the amplitudes of the stimuli are sufficiently large (Geisler & Albrecht, 1995).

Relationship to behavioral discrimination performance

In this study, the descriptive function method was used to measure discrimination performance for a large population of cells in monkey and cat. The results provide a unique opportunity to compare cortical neuron responses with behavioral psychophysics. To this end, we compared neural and behavioral discrimination performance along the dimensions of contrast and spatial frequency, utilizing two different methods. In the first method, we determined the best performance of each cell individually and then plotted these individual performances against published measures of behavioral performance. In the second method, we determined the performance of the entire population of cells as a whole, assuming optimal pooling of neural responses, and then plotted the population performance against the behavioral performance.

The solid symbols in Fig. 10A show the best contrast discrimination performance for all of the neurons recorded in monkey. Each point represents the minimum of the contrast discrimination function (i.e. the minimum contrast threshold, and the base contrast where the minimum occurred, see arrow in Fig. 5A). As can be seen, these best performance points are widely scattered at each base contrast, and the best base contrasts span most of the contrast axis. All points that fall on the vertical axis at 0% contrast represent cells where the best discrimination threshold was the absolute threshold; all other points represent cells where the threshold initially decreased with contrast (i.e. they showed the "dipper" effect). It is interesting to note that as contrast increases the thresholds of the most sensitive cells increase monotonically. Specifically, those cells with the lowest thresholds are bounded by an envelope which is approximately constant at lower contrasts and linear (in log–log coordinates) at higher contrasts.[7]

The open symbols in Fig. 10A plot behavioral contrast discrimination functions for monkeys (Kiper & Kiorpes, 1994) and humans (Bradley & Ohzawa, 1986; Legge & Kersten, 1987; Seay & Geisler, 1995). As contrast increases, threshold initially decreases by a small amount (the "dipper" effect), and then increases linearly with a slope of 0.6 to 1.0 (i.e. as a power function with an exponent between 0.6 and 1.0). As can be seen, the general shape of the behavioral functions is similar to the shape of the envelope of the most sensitive cells. However, the behavioral thresholds are somewhat lower.

The envelope of the most sensitive cells provides one method for comparing cortical and behavioral performance. This comparison is appropriate if behavioral discrimination at each contrast is determined by the most sensitive neuron at each contrast. On the other hand, behavioral discrimination at each contrast could be determined by combining the responses of all neurons. Although it is clear that the most sensitive neurons at each contrast will contribute the most information, there is nevertheless additional information in the responses of the other neurons.

The solid curve in Fig. 10B shows the *population contrast discrimination function*; this curve is the result of pooling all of the discrimination information from all of the cells in an optimal fashion. Each point on this curve was obtained as follows. First, the measured contrast response function for each neuron was used to find the mean responses and standard deviations to the base contrast and to the base-plus-increment. Second, these means and standard deviations were used to obtain the signal-to-noise ratio (the d') for each cell [see eqn. (2)]. Third, the d' summation formula from Bayesian decision theory was used to find the signal-to-noise ratio for the entire population [see eqn. (3)]. Fourth, the threshold was obtained by varying the increment contrast until the population d' equaled 1.0 (75% correct). Finally, the population function was shifted vertically, using an efficiency parameter, to allow comparison of the shapes (see Methods).

The open symbols in Fig. 10B are the same behavioral data shown in Fig. 10A. As can be seen, the shape of the population discrimination function is remarkably similar to the shapes of the behavioral discrimination functions.

The two methods of comparing behavioral and neural performance (illustrated in Figs. 10A and 10B) represent two extremes in psychophysical linking hypotheses. At one extreme, behavioral performance is modeled by the envelope of the most sensitive neurons (Barlow, 1972, 1995; De Valois et al., 1967; Talbot et al., 1968). At the other extreme, behavioral performance is modeled by optimally combining the responses of all neurons that might contribute rele-

[7] It is important to consider the possibility that the variations in sensitivity with contrast could be the result of correlations between the preferred spatial frequency and the parameters of the contrast response function. For example, the high spatial frequency cells in this population might account for the increase in the thresholds at higher base contrasts if there were a correlation between preferred spatial frequency and the semisaturation constant. To evaluate this possibility, we measured the correlations between preferred spatial frequency and the parameters of the contrast response function. No significant correlations were found. Further, when the analysis was restricted to those cells below 10 cpd, the overall trends were the same.

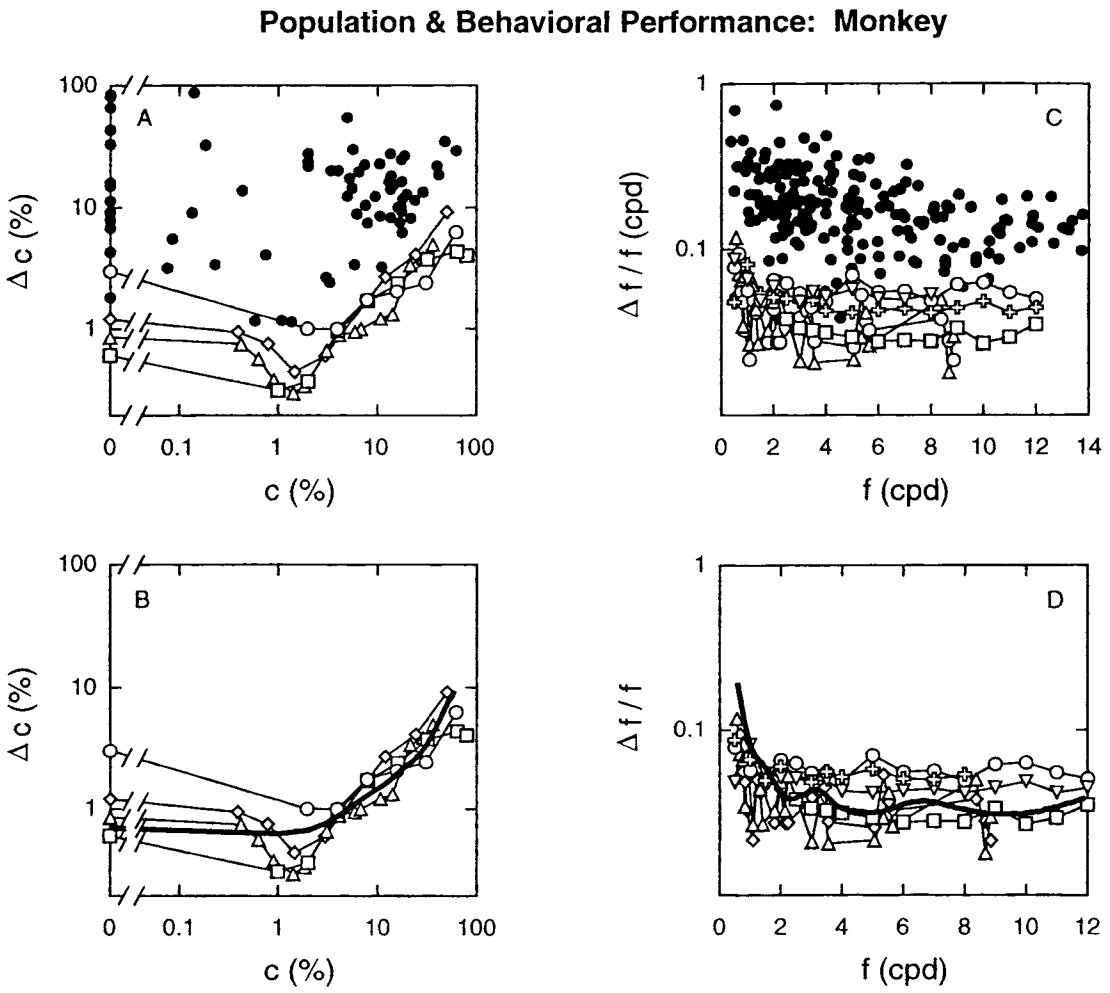

Fig. 10. Cortical performance and behavioral performance: Monkey. Comparison of the discrimination performance of monkey visual cortex neurons with the behavioral discrimination performance of monkeys and humans, using two different methods: the envelope of the most sensitive cells (upper panels), and optimal pooling (lower panels). In A, the solid symbols plot the best contrast discrimination thresholds for each neuron (i.e. the minimum Δc, and the location along the contrast axis where the minimum occurred); the open symbols plot behavioral contrast discrimination thresholds for monkeys (○, Kiper & Kiorpes, 1994) and humans (◇, Bradley & Ohzawa, 1986; △, Legge & Kersten, 1987; □, Seay & Geisler, 1995). In B, the solid line plots the overall contrast discrimination performance for the population of cells as a whole, assuming that the discrimination information from all cells is pooled in an optimal fashion (using Bayesian decision theory); the open symbols replot the behavioral contrast discrimination thresholds. The neural population curve was shifted vertically (by varying an efficiency parameter, $\epsilon_{100} = 0.5$) to allow comparison of the shapes. In C, the solid symbols plot the best spatial-frequency discrimination thresholds for the neurons; the open symbols plot spatial-frequency discrimination for humans (○, □, ▽, ✣, Richter & Yager, 1984; △, ◇, Blake et al. 1986). In D, the solid line plots spatial-frequency discrimination for the population of cells, assuming optimal pooling, and the open symbols replot behavioral spatial-frequency discrimination. Again, the neural population curve was shifted vertically (using the same value of the efficiency parameter, $\epsilon_{100} = 0.5$).

vant information (Watson, 1983; Geisler, 1984, 1989). The results in Figs. 10A and 10B show that both models account for the data about equally well. Presumably, models between the two extremes could also account for the data reasonably well. Thus, it may prove difficult to distinguish between these different models (see Parker & Newsome, 1998).

On the other hand, it is important to note that the contrast discrimination functions of individual cortical neurons are quite different from the behavioral discrimination functions, in that the most sensitive neurons only cover a limited range of contrasts (e.g. see Figs. 5, 9, and Fig. 16 in Appendix C). Thus, it seems likely that behavioral contrast discrimination functions reflect the combined sensitivity of a heterogeneous population of neurons (Albrecht & Hamilton, 1982; Barlow et al., 1987).

Interestingly, there is little or no indication of a dipper in the population performance even though most individual cells have a dipper. The heterogeneity of individual cortical cells could perhaps explain this result: the dippers for the individual cells are widely distributed along the contrast axis. Thus, our results suggest that the dipper effect observed in psychophysical experiments is probably not the result of the response properties of neurons in the primary visual cortex. On the other hand, the results do not exclude existing psychophysical explanations for the dipper effect, including neural thresholds (Foley & Legge, 1981), accelerating nonlinearities (Legge & Foley, 1980), and/or signal uncertainty (Pelli, 1985), because these mechanisms could operate at neural stages beyond V1.

Figs. 10C and 10D show a comparison of behavioral and neural spatial-frequency discrimination. The solid symbols in Fig. 10C

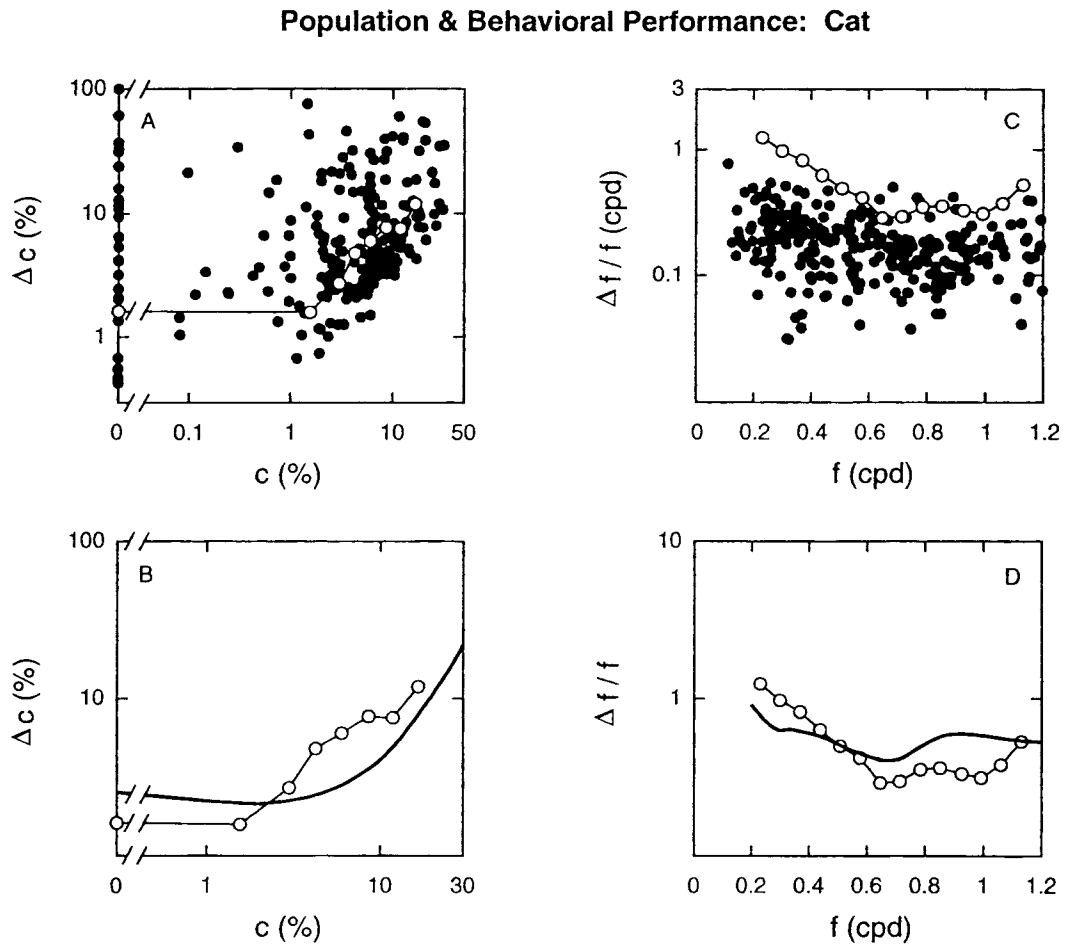

Fig. 11. Cortical performance and behavioral performance: Cat. Comparison of the discrimination performance of cat visual cortex neurons with the behavioral discrimination performance of cats, using the two different methods (most sensitive cells, and optimal pooling). The conventions are the same as those described in Fig. 10. In A, the solid symbols plot the best contrast discrimination thresholds for each neuron and the open symbols plot behavioral contrast discrimination thresholds (Blake & Petrakis, 1984). In B, the solid line plots the contrast discrimination for the population of cells (using optimal pooling with an efficiency parameter, $\epsilon_{100} = 0.04$), and the open symbols replot the behavioral contrast discrimination thresholds. In C, the solid symbols plot the best spatial-frequency discrimination thresholds for the neurons; the open symbols plot the spatial-frequency discrimination thresholds in cats (Blake et al., 1986). In D, the solid line plots spatial-frequency discrimination for the population of cells ($\epsilon_{100} = 0.006$) and the open symbols replot behavioral spatial-frequency discrimination.

plot the best spatial-frequency discrimination performance for each individual neuron; the open symbols plot representative spatial-frequency discrimination functions for humans (Richter & Yager, 1984; Blake et al., 1986). Again, the overall shape of the behavioral functions is similar to the shape of the envelope of the most sensitive cells (i.e. the shape is approximately flat with a slight rise at low frequencies); however, the behavioral thresholds are generally lower. The solid curve in Fig. 10D shows the neural population function for spatial-frequency discrimination along with the behavioral functions. As can be seen, the shapes of the neural and behavioral functions are quite similar.

Fig. 11 shows a similar analysis for cat. Fig. 11A compares behavioral performance (Blake & Petrakis, 1984) with the best performance of individual neurons along the dimension of contrast. As in monkey, the envelope of the most sensitive cells parallels the behavioral discrimination function except in this case the neural envelope is below the behavioral function rather than above. Fig. 11B compares the behavioral performance with the population contrast discrimination function. Fig. 11C compares the behavioral performance (Blake et al., 1986) with the best performance of individual neurons along the dimension of spatial frequency. Again, the envelope of the most sensitive cells parallels the behavioral function but now the envelope is far below the behavioral function. Finally, Fig. 11D compares the behavioral performance with the population spatial-frequency discrimination function.

Interestingly, for the monkey, the value of the efficiency parameter needed to bring the population discrimination functions into agreement with the behavioral discrimination functions (Figs. 10B and 10D) was the same for both contrast and spatial frequency ($\epsilon_{100} = 0.5$). This suggests that the neural information for these two dimensions is being utilized with equivalent efficiency by subsequent brain mechanisms. In the cat, the efficiencies are considerably lower than in the monkey and are quite different for contrast ($\epsilon_{100} = 0.04$) and spatial frequency ($\epsilon_{100} = 0.006$). This relatively poor utilization of neural information for spatial-frequency discrimination is consistent with the analysis of cat behavioral performance by Blake et al. (1986).

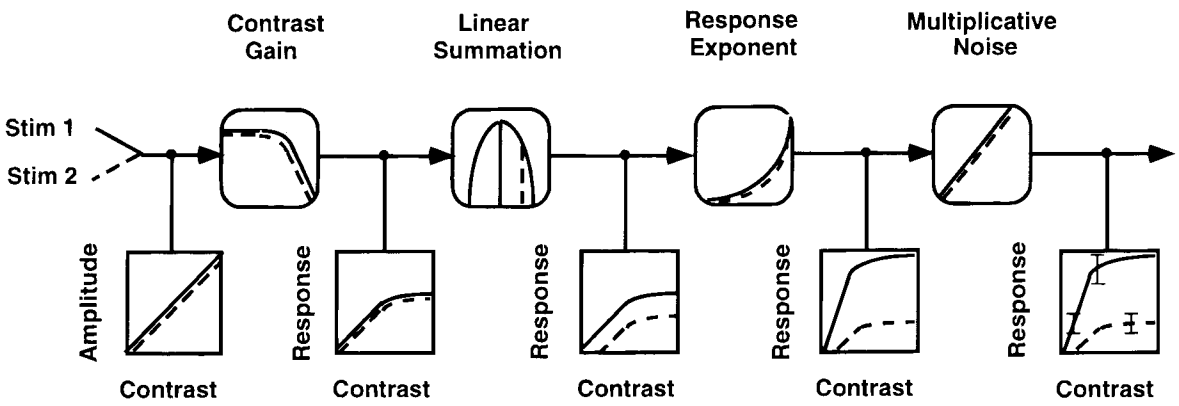

Fig. 12. Contrast-gain exponent model: CGE filter. Components of the contrast-gain exponent model (upper boxes) and the effect of each component (lower boxes) on the contrast response function for an optimal (solid curves) and nonoptimal (dashed curves) stimulus, plotted in log–log coordinates. The contrast-gain mechanism is a fast-acting gain control that scales the input amplitude by a factor which decreases with increasing contrast, thereby producing saturation at higher contrasts; it is set by the local mean contrast, not the response rate, and thus the optimal and the nonoptimal stimulus are attenuated equally. The linear summation mechanism is a filter which gives the cell its fundamental selectivities; it causes the nonoptimal stimulus to be attenuated more than the optimal stimulus. The response exponent mechanism takes the output of the linear filter to an exponent greater than 1.0, thus increasing the difference in the responses to the optimal and nonoptimal stimuli. The multiplicative noise mechanism introduces a response variance that is proportional to the mean response rate; thus, when the mean responses to an optimal and a nonoptimal stimulus are equal, their variances will be equal.

Neuron sampling models for behavioral discrimination

The results from this report (and many others: see references in the Introduction) make it clear that cortical neurons are heterogeneous in their stimulus–response relationships; they are continuously variable in preferred spatial frequency, preferred orientation, spatial-frequency bandwidth, orientation bandwidth, half-saturation contrast, response exponent, maximum response, variance constant, etc. (see the population statistics in Table 2). Despite this heterogeneity, it is possible to summarize the stimulus–response relationships of each cell using a small set of relatively simple descriptive functions. For example, for the population of cells in this study, the descriptive functions in Table 1 were able to account for 90% of the variability in both the mean response and the standard deviation.

Many quantitative psychophysical models of visual discrimination consist of a discrete set of tuned channels, together with a single contrast nonlinearity. Given the continuous heterogeneity of cortical cells along the various dimensions measured in this and other studies, as well as the continuous heterogeneity of the nonlinearities, it seems reasonable to consider an alternate class of models. A more parsimonious approach is to represent the early visual system as a large collection of neurons whose tuning characteristics, nonlinear response characteristics, and noise characteristics are randomly sampled from probability distributions. In this framework, the probability distributions of the cortical cell properties take the place of the discrete channels.

To construct such models, *neuron sampling models*, it is necessary to determine the probability distributions for the cell properties (e.g. it is necessary to determine the probability distributions for each of the descriptive parameters in Table 1). One approach is to estimate the parameters of the distributions by fitting psychophysical data. The other approach is to take the parameters from published measurements of cortical cell properties (Seay et al., 1990). With this latter approach, there are very few parameters to be estimated from the psychophysical data.

The fact that the neural and the behavioral performances are in good agreement for the random sample of cells in this study (see Figs. 10 and 11) suggests that neuron sampling models may prove to be valuable. Further, setting aside the notion of discrete channels (or pathways), and placing the emphasis on the distributions of cell properties, might lead to a useful shift in theoretical perspective.

Implications for pattern recognition

The results summarized in Figs. 7 and 8 show that when a cell produces a response near the maximum, subsequent brain mechanisms can be certain that the stimulus lies within a narrow range. In other words, when the responses are large, single cortical cells can signal, with reasonable reliability, the presence of specific stimulus attributes. This suggests that individual cortical cells might play a more important role in pattern recognition than expected, given their broad tuning, low response rate, and high response variability. However, this suggestion is based upon the measured responses to sine waves presented one at a time. Sine waves provide a good match to the receptive field and hence produce robust responses. Natural images, on the other hand, contain a broad range of sine-wave components. Most of these components do not match the receptive field and may actually cause response suppression (e.g. as in contrast normalization). Thus, it is not at all certain that natural images would produce responses in single neurons that are large enough to reliably signal specific stimulus attributes (Gallant et al., 1995).

To evaluate how individual cortical neurons would respond to complex natural images, we utilized a recent model which is consistent with the descriptive functions given in Table 1. It has been shown (Albrecht & Geisler, 1991, 1994; Heeger, 1991, 1992*a*,*b*; Geisler & Albrecht, 1995) that the responses of cortical neurons are adequately described by a model consisting of four components: (1) a linear filter which establishes the neuron's stimulus selectivity, (2) a contrast normalization mechanism which main-

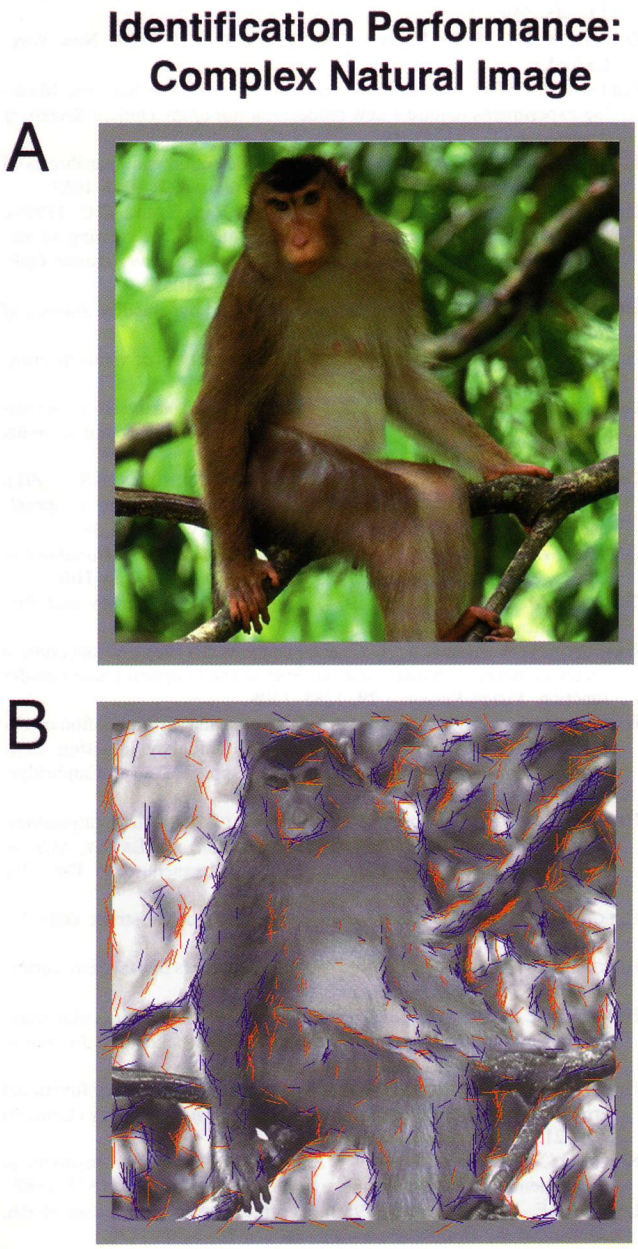

Fig. 13. Identification performance: Complex natural image. Estimated responses of single visual cortex neurons to a complex natural image. (A) An image of a macaque monkey in a natural setting. The RMS contrast of the image was 0.5. (B) An achromatic reduced-contrast version of the image, overlayed with colored line segments indicating the neurons that responded at greater than 50% of their maximum response. The position, length, orientation, and color of the line segments indicate the position, size, orientation, and polarity of each neuron's receptive field. (Polarity refers to the spatial sequence of light–dark preference of the receptive field: blue = light-to-dark, red = dark-to-light.) There is a reasonable correspondence between the properties of the line segments and the local image features (e.g. note the series of blue line segments along the light-to-dark boundary at the edge of the right arm). The responses of the neurons were determined assuming that the size of the image was 4 deg × 4 deg, and fixation was at the center. The reader can approximate this stimulus situation by fixating the center of the image from a distance of approximately 1 m (arm's length). (Photograph from the Frans Lanting/National Geographic Image Collection.)

tains selectivity in spite of response saturation, (3) an expansive nonlinearity which enhances selectivity, and (4) a noise source which makes the variance proportional to the mean. A schematic of this model, the *contrast-gain exponent model*, is shown in Fig. 12, and the formal mathematical description is given in Appendix D. It is important to emphasize that this model is meant to provide a functional description of single cortical neuron responses and should not be taken as a hypothesis about the detailed anatomy and physiology. For example, contrast normalization is probably occurring at many levels, starting in the retina (Shapley & Victor, 1979; Sclar, 1987; Albrecht & Geisler, 1991; Albrecht, 1995).

The contrast-gain exponent model was generalized to a population of cells covering a 4 deg × 4 deg field in a fashion consistent with the concept of cortical magnification, and was then used to estimate the response of the visual cortex to a 4 deg × 4 deg natural image.

Fig. 13B illustrates how such a population of cortical cells would respond to the complex natural image in Fig. 13A. Rather than show the responses of all cells, we show the responses of the cells with a critical frequency approximately 1.5 octaves below the cutoff of the monkey's CSF at each eccentricity (e.g. 16 cpd in the center of the fovea and 4.4 cpd at 2-deg eccentricity), with fixation at the center of the image. Each line segment drawn on the image indicates a single neuron that produced at least 50% of the maximum response. The length and orientation of the line segment indicates the length and orientation of the receptive field. The red line segments indicate "edge-like" receptive fields sensitive to a light–dark transition in one direction (90-deg phase), and the blue line segments indicate "edge-like" receptive fields sensitive to a light–dark transition in the opposite direction (270-deg phase). Comparison of Figs. 13A and 13B shows that the orientation, location, and phase of each line segment corresponds reasonably well to the orientation, location, and phase of specific local contours in the complex image. For example, note the blue line segments which follow the light to dark edge along the monkey's right arm, and the red line segments which follow the dark to light edge along the monkey's right cheek. Thus, even though cortical cells are broadly tuned and noisy, when they produce a large response (5–10 spikes in 200 ms) they transmit accurate information about local image features.

The cells that produced weak responses (10%–50% of the maximum) are not shown in Fig. 13. These cells provided relatively poor pattern-recognition information; the orientation and location of their receptive fields were generally poorly correlated with the orientation and location of the contours in the image.

Note that there are many regions in the image where there are no line segments. There are a number of reasons for this. First, for computational reasons, the number of cells we used was approximately 1.5 orders of magnitude less than the actual number of cortical cells. Second, the density of the cells decreases with eccentricity. Third, for ease of viewing, we show only the responses of one size receptive field at each eccentricity. Fourth, even within the subpopulation, only 0.4% of the cells produced a response that exceeded the 50% maximum-response criterion.[8] If the number of cells in the simulation was increased to that in the primary visual cortex, then the number of cells producing responses greater than

[8] This occurs because (1) the image does not contain all orientations and phases, at all locations; and (2) even when an image contour is at the correct location, orientation, and phase, the shape of the contour may not perfectly match the receptive field.

50% of the maximum response would form a fairly dense representation of the contours in the image.

Based upon the measurements of the 95% confidence regions, and the results in Fig. 13, it seems reasonable to conclude that local image contours can be represented by the activity of individual neurons in the primary visual cortex. On the other hand, even though the 95% confidence regions are quite small, they are probably not small enough to account for the very precise local representation achieved by the visual system as a whole. Thus, it seems quite possible that rudimentary local feature recognition is accomplished by the level of V1 and then refined in subsequent cortical stages. Hopefully, the methods developed here can be used to measure visual performance in these other more advanced regions of the cortex.

Acknowledgments

This research was supported by the Air Force Office of Scientific Research (F49620-93-1-0307), National Eye Institute (EY02688), and The University of Texas.

References

ALBRECHT, D.G. (1995). Visual cortex neurons in monkey and cat: Effect of contrast on the spatial and temporal phase transfer functions. *Visual Neuroscience* **12**, 1191–1210.

ALBRECHT, D.G., FARRAR, S.B. & HAMILTON, D.B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *Journal of Physiology* **347**, 713–739.

ALBRECHT, D.G. & GEISLER, W.S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience* **7**, 531–546.

ALBRECHT, D.G. & GEISLER, W.S. (1994). Visual cortex neurons in monkey and cat: Contrast response nonlinearities and stimulus selectivity. In *Computational Vision Based on Neurobiology*, ed. LAWTON, T., Vol. 2054, pp. 12–31. Bellingham, Washington: SPIE.

ALBRECHT, D.G. & HAMILTON, D.B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology* **48**, 217–237.

BARLOW, H.B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception* **1**, 371–394.

BARLOW, H.B. (1995). The neuron doctrine in perception. In *The Cognitive Neurosciences*, ed. GAZZANIGA, M.S., pp. 415–435. Cambridge, Massachusetts: The MIT Press.

BARLOW, H.B., KAUSHAL, T.P., HAWKEN, M. & PARKER, A.J. (1987). Human contrast discrimination and the threshold of cortical neurons. *Journal of the Optical Society of America A* **4**, 2366–2371.

BLAKE, R., HOLOPIGIAN, K. & WILSON, H.R. (1986). Spatial-frequency discrimination in cats. *Journal of the Optical Society of America A* **3**, 1443–1449.

BLAKE, R. & PETRAKIS, I. (1984). Contrast discrimination in the cat. *Behavioral Brain Research* **12**, 155–162.

BRADLEY, A. & OHZAWA, I. (1986). A comparison of contrast detection and discrimination. *Vision Research* **26**, 991–997.

BRADLEY, A., SKOTTUN, B.C., OHZAWA, I., SCLAR, G. & FREEMAN, R.D. (1985). Neurophysiological evaluation of the differential response model for orientation and spatial-frequency discrimination. *Journal of the Optical Society of America A* **2**, 1607–1610.

CAMPBELL, F.W., COOPER, G.F. & ENROTH-CUGELL, C. (1968). The spatial selectivity of visual cells of the cat. *Journal of Physiology* **203**, 223–235.

CAMPBELL, F.W., COOPER, G.F., ROBSON, J.G. & SACHS, M.B. (1969). The spatial selectivity of visual cells of the cat and the squirrel monkey. *Journal of Physiology* **204**, 120–121P.

CHANDLER, J.P. (1969). STEPIT—Finds local minima of a smooth function of several parameters. *Behavioral Science* **14**, 81–82.

CURCIO, C.A. & ALLEN, K.A. (1990). Topography of ganglion cells in human retina. *Journal of Comparative Neurology* **300**, 5–25.

DE VALOIS, R.L., ABRAMOV, I. & MEAD, W.R. (1967). Single cell analysis of wavelength discrimination at the lateral geniculate nucleus in the macaque. *Journal of Neurophysiology* **30**, 415–433.

DE VALOIS, R.L., ALBRECHT, D.G. & THORELL, L.G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research* **22**, 545–559.

DE VALOIS, R.L. & DE VALOIS, K.K. (1988). *Spatial Vision*. New York: Oxford.

FOLEY, J.M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A* **11**, 1710–1719.

FOLEY, J.M. & LEGGE, G.E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research* **21**, 1041–1053.

GALLANT, J.L., CONNER, C.E., DRURY, H. & VAN ESSEN, D.C. (1995). Neural responses in monkey visual cortex during free viewing of natural scenes: Mechanisms of response suppression. *Investigative Ophthalmology and Visual Science* (Suppl.) **36/4**, S1052.

GEISLER, W.S. (1984). Physical limits of acuity and hyperacuity. *Journal of the Optical Society of America A* **1**, 775–782.

GEISLER, W.S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review* **96**, 267–314.

GEISLER, W.S. & ALBRECHT, D.G. (1995). Bayesian analysis of identification in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. *Vision Research* **35**, 2723–2730.

GEISLER, W.S., ALBRECHT, D.G., SALVI, R.J. & SAUNDERS, S.S. (1991). Discrimination performance of single neurons: Rate and temporal-pattern information. *Journal of Neurophysiology* **66**, 334–361.

GEISLER, W.S. & BANKS, M.S. (1995). Visual performance. In *Handbook of Optics*, ed. BASS, M., pp. 25.1–25.55. New York: McGraw-Hill.

GREEN, D.M. & SWETS, J.A. (1974). *Signal Detection Theory and Psychophysics*. New York: Krieger.

HAMILTON, D.B., ALBRECHT, D.G. & GEISLER, W.S. (1989). Visual cortical receptive fields in monkey and cat: Spatial and temporal phase transfer function. *Vision Research* **29**, 1285–1308.

HAWKEN, M.J. & PARKER, A.J. (1990). Detection and discrimination mechanisms in the striate cortex of the Old-World monkey. In *Vision: Coding and Efficiency*, ed. BLAKEMORE, C., pp. 103–116. Cambridge, Massachusetts: Cambridge University Press.

HEEGER, D.J. (1991). Computational model of cat striate physiology. In *Computational Model of Visual Perception*, ed. LANDY, M.S. & MOVSHON A., pp. 119–133. Cambridge, Massachusetts: The MIT press.

HEEGER, D.J. (1992*a*). Half-squaring in responses of cat striate cells. *Visual Neuroscience* **9**, 427–443.

HEEGER, D.J. (1992*b*). Normalization of cell responses in cat striate cortex. *Visual Neuroscience* **9**, 191–197.

HUBEL, D.H. & WIESEL, T.N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology* (London) **160**, 106–154.

HUBEL, D.H., & WIESEL, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* (London) **195**, 215–243.

KIPER, D.C. & KIORPES, L. (1994). Suprathreshold contrast sensitivity in experimentally strabismic monkeys. *Vision Research* **34**, 1575–1583.

LEGGE, G.E. & FOLEY, J.M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America* **70**, 1458–1470.

LEGGE, G.E. & KERSTEN, D. (1987). Contrast discrimination in peripheral vision. *Journal of the Optical Society of America A* **4**, 1594–1598.

MOOD, A.M. & GAYBILL, F.A. (1963). *Introduction to the Theory of Statistics* (second ed.). San Francisco, California: McGraw-Hill.

ORBAN, G.A. (1984). *Neuronal Operations in the Visual Cortex*. New York: Springer-Verlag.

PALMER, L.A., JONES, J.P. & STEPNOSKI, R.A. (1991). Striate receptive fields as linear filters: Characterization in two dimensions of space. In *The Neural Basis of Visual Function*, ed. LEVENTHAL, A.G., pp. 246–265. Boston, Massachusetts: CRC Press.

PARKER, A.J., & HAWKEN, M.J. (1985). The capabilities of cortical cells in spatial-resolution tasks. *Journal of the Optical Society of America A* **2**, 1101–1114.

PARKER, A.J. & NEWSOME, W.T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience* **21** (in press).

PELLI, D.G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A* **2**, 1508–1532.

RICHTER, E.S. & YAGER, D. (1984). Spatial-frequency difference thresholds for central and peripheral viewing. *Journal of the Optical Society of America A* **1**, 1136–1139.

ROBSON, J.G. (1975). Receptive fields: Spatial and intensive representation of the the visual image. In *Handbook of Perception Vol. 5: Vision*, ed. CARTERETTE, E.C. & FRIEDMAN, M.P., pp. 81–112. New York: Academic Press.

SCLAR, G. (1987). Expression of "retinal" contrast gain control by neurons of the cat's lateral geniculate nucleus. *Experimental Brain Research* **66**, 589–596.

SCLAR, G. & FREEMAN, R.D. (1982). Orientation selectivity in the cat's striate cortex is invariant with contrast. *Experimental Brain Research* **46**, 457–461.

SEAY, C.A. & GEISLER, W.S. (1995). Signal detection analysis of contrast matching and discrimination. *Investigative Ophthalmology & Visual Science Supplement* **36/7**, S903.

SEAY, C.A., GEISLER, W.S. & ALBRECHT, D.G. (1996). A neuron sampling model of spatial vision. *Investigative Ophthalmology & Visual Science Supplement* **37/3**, S233.

SHAPLEY, R.M. & LENNIE, P. (1985). Spatial frequency analysis in the visual system. *Annual Review of Neuroscience* **8**, 547–583.

SHAPLEY, R.M. & VICTOR, J.D. (1979). The contrast gain control of the cat retina. *Vision Research* **19**, 431–434.

SNOWDEN, R.J., TREUE, S. & ANDERSEN, R.A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research* **88**, 389–400.

SOFTKY, W.R. & KOCH, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *Journal of Neuroscience* **13**, 334–350.

TALBOT, W.P., DARIAN-SMITH, I., KORNHUBER, H.H. & MOUNTCASTLE, V.B. (1968). The sense of flutter-vibration: Comparison of human capacity with response patterns of mechano-receptive afferents from the monkey hand. *Journal of Neurophysiology* **31**, 301–334.

TOLHURST, D.J., MOVSHON, J.A. & DEAN, A.F. (1983). The statistical reliability of signals in single neurons in the cat and monkey visual cortex. *Vision Research* **23**, 775–785.

VOGELS, R., SPILEERS, W. & ORBAN, G.A. (1989). The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research* **77**, 432–436.

WATSON, A.B. (1983). Detection and recognition of simple spatial forms. In *Physical and Biological Processing of Images*, ed. BRADDICK, O.J. & SLEIGH, A.C., pp. 110–114. Berlin: Springer-Verlag.

WATSON, A.B. (1986). Temporal sensitivity. In *Handbook of Perception and Human Performance*, ed. BOFF, K.R., KAUFMAN, L. & THOMAS, J.P., pp. 1–43. New York: John Wiley and Sons.

Appendix A: Maximum-likelihood fitting of descriptive functions

This appendix describes the maximum-likelihood procedure we used for fitting the descriptive functions to the response means and response variances. Such maximum-likelihood procedures are known to yield parameter estimates that are optimal in a variety of ways (Mood & Gaybill, 1963).

On each trial of an experiment a particular stimulus is presented; let $\mathbf{x}_j$ represent the jth stimulus. A particular stimulus is generally presented more than once. We represent the number of spikes observed in the ith presentation of stimulus $\mathbf{x}_j$ as $Z_i(\mathbf{x}_j)$. For purposes of parameter estimation, we assume that the $Z_i(\mathbf{x}_j)$ are statistically independent, and are approximately normally distributed with a mean of $r(\mathbf{x}_j)$ and a variance $\sigma^2(\mathbf{x}_j)$:

$$\phi_j(z) = \frac{1}{\sqrt{2\pi\sigma^2(\mathbf{x}_j)}} e^{-\frac{1}{2}\frac{[z-r(\mathbf{x}_j)]^2}{\sigma^2(\mathbf{x}_j)}} \tag{A1}$$

where $r(\mathbf{x})$ and $\sigma^2(\mathbf{x})$ are the descriptive functions for the mean and variance, respectively. It is the parameters in these descriptive functions that were estimated in the fitting process.

The probability or likelihood (L) of the measured responses is simply the product of the probabilities of each of the measured responses:

$$L = \prod_{j=1}^{m}\prod_{i=1}^{n} \phi_j[Z_i(\mathbf{x}_j)] = \prod_{j=1}^{m}\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2(\mathbf{x}_j)}} e^{-\frac{1}{2}\frac{[Z_i(\mathbf{x}_j)-r(\mathbf{x}_j)]^2}{\sigma^2(\mathbf{x}_j)}} \tag{A2}$$

Maximum-likelihood estimates are obtained by maximizing L with respect to the unknown parameters in the equations $r(\mathbf{x})$ and $\sigma^2(\mathbf{x})$. To improve computational precision and speed, it is usually preferable to maximize the logarithm of the likelihood, or equivalently to minimize $-2\ln L$,

$$-2\ln L = nm\ln(2\pi) + \sum_{j=1}^{m}\left\{ n\ln[\sigma^2(\mathbf{x}_j)] + \frac{1}{\sigma^2(\mathbf{x}_j)}\sum_{i=1}^{n}[Z_i(\mathbf{x}_j) - r(\mathbf{x}_j)]^2 \right\} \tag{A3}$$

Minimizing $-2\ln L$ is equivalent to maximizing the likelihood because $-2\ln L$ is a monotonically decreasing function of L.

Eqn. (A3) can be used to estimate parameters when the data consist of spike counts for each trial. However, for some of our experiments, only the mean and the standard deviation of the spike count were saved. Fortunately, saving the individual spike counts is unnecessary; in fact, identical parameter estimates are obtained by minimizing the following equation, which can be obtained from eqn. (A3) with some algebra:

$$-2\ln L = nm\ln(2\pi) + \sum_{j=1}^{m}\left\{ n\ln[\sigma^2(\mathbf{x}_j)] + \frac{n[S^2(\mathbf{x}_j) + \bar{Z}^2(\mathbf{x}_j) - 2r(\mathbf{x}_j)\bar{Z}(\mathbf{x}_j) + r^2(\mathbf{x}_j)]}{\sigma^2(\mathbf{x}_j)} \right\} \tag{A4}$$

In this equation, $\bar{Z}(\mathbf{x}_j)$ and $S(\mathbf{x}_j)$ represent the observed mean and standard deviation of the spike counts measured for stimulus level j.

Descriptive functions were fitted to the measured spike data by minimizing eqn. (A4), using the program STEPIT (Chandler, 1969). In general, local minima proved not to be a significant problem because the parameters of the descriptive functions are relatively orthogonal, and because the parameters were initialized to values typical of cortical cells. Although the fitting process was automated, each best fit was plotted against the data and visually inspected to ensure that both the fit and the final parameters were reasonable.

Appendix B: The relationship between the response mean and the response variance

Previous investigators (e.g. Tolhurst et al., 1983; Vogels et al., 1989; Snowden et al., 1992; Softky & Koch, 1993) have described the relationship between the mean and variance of cortical neuron responses with a two-parameter power function:

$$\sigma^2 = Kr^{\nu} \tag{B1}$$

where K is a proportionality constant, and ν an exponent. The reported average values of K typically range from 1.2 to 1.5 and the values of ν range from 1.0 to 1.2. Initially, we used eqn. (B1) to describe the mean-variance relationship; however, we found that the two parameters, K and ν, were highly correlated, and this led us to also consider the simpler formula:

$$\sigma^2 = Kr \tag{B2}$$

This appendix describes a detailed analysis comparing the two descriptive functions. We show that the simple proportionality function [eqn. (B2)] works quite well, and that there is little to be gained by using the more complex power function [eqn. (B1)], at least for the purposes of characterizing tuning characteristics, and measuring detection, discrimination, and identification performance.

Standard statistical techniques (Mood & Gaybill, 1963) were used to rigorously determine whether the power function provided a significantly better fit than the proportionality function. First, maximum-likelihood estimates of the proportionality constant and exponent in eqn. (B1) were obtained for each neuron. Then, the maximum-likelihood estimates of the proportionality constant in eqn. (B2) were obtained for each neuron. To provide the fairest possible comparison of the two functions, we left the means completely unconstrained (i.e. we modeled each mean as a free

parameter). This ensured that the estimates of K and ν were not influenced by the choice of descriptive formulas for the means. Using the results of the maximum-likelihood procedure, a χ^2 test was applied to determine whether the power function provided a significantly better fit than the proportionality function.

To obtain the unconstrained maximum-likelihood estimates of K and ν, the formulas for the variance [eqns. (B1) and (B2)] were substituted into eqn. (A4). Thus, the likelihood given the two-parameter power function is

$$-2\ln L = nm\ln(2\pi) + \sum_{j=1}^{m}\left\{n\ln(Kr_j^{\nu}) + \frac{n[S^2(\mathbf{x}_j) + \bar{R}^2(\mathbf{x}_j) - 2r_j\bar{R}(\mathbf{x}_j) + r_j^2]}{Kr_j^{\nu}}\right\} \tag{B3}$$

and the likelihood given the one-parameter proportionality function is

$$-2\ln L_1 = nm\ln(2\pi) + \sum_{j=1}^{m}\left\{n\ln(Kr_j) + \frac{n[S^2(\mathbf{x}_j) + \bar{R}^2(\mathbf{x}_j) - 2r_j\bar{R}(\mathbf{x}_j) + r_j^2]}{Kr_j}\right\} \tag{B4}$$

Using the minimization program STEPIT (Chandler, 1969), we obtained simultaneous maximum-likelihood estimates of the parameters of the power-function (K, ν) and of the means (the r_js). In a separate analysis, we obtained simultaneous estimates of the parameter of the proportionality function (K) and of the means (the r_js).

A useful property of maximum-likelihood estimation is that the results can be used to test whether one model provides a significantly better fit to the data than another model. Specifically, the difference in the minimum value of $-2\ln L$ (obtained with the proportionality and power functions) is asymptotically chi-square distributed with 1 deg of freedom, under the null hypothesis that the functions are equivalent (Mood & Gaybill, 1963):

$$\chi^2_{(1)} = \min\{-2\ln L_1\} - \min\{-2\ln L_2\} \tag{B5}$$

If the computed value of $\chi^2_{(1)}$ is significantly greater than 1.0, we can conclude that the power function provides a better fit.

Fig. 14 shows the frequency histograms of the $\chi^2_{(1)}$ values for all the cells. The first bin (for cat and monkey) represents the number of cells which differed at the 0.05 level. For approximately 19% of the monkey cells and 21% of the cat cells, the power function was significantly better; for the remaining cells it was not. If the two functions were equivalent, one would expect (by chance) 5% of the cells to be fit significantly better by the power function. Thus, the power function fitted some of the cells better

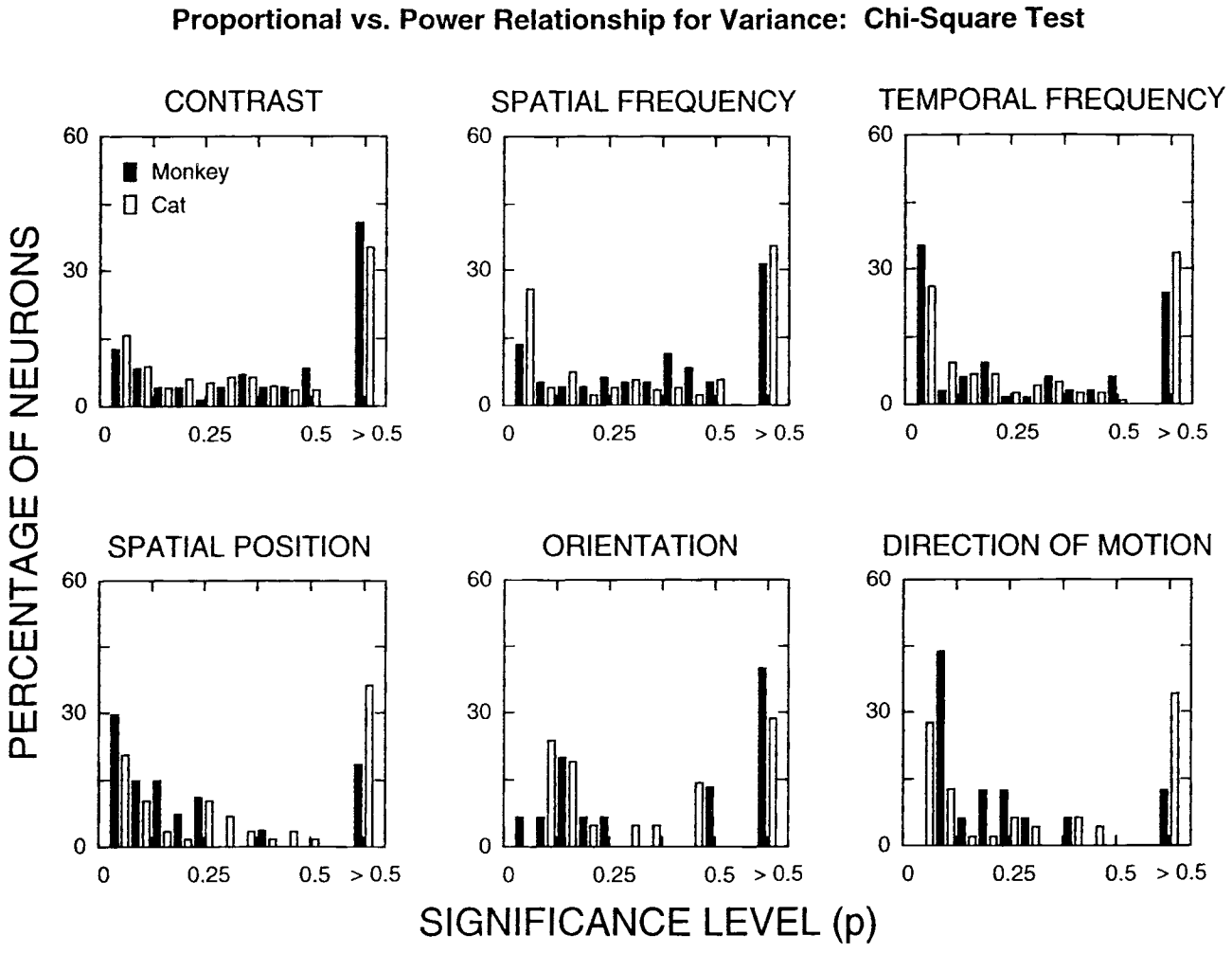

Fig. 14. Chi-square test of mean-variance relationship. Histograms showing the significance levels for a chi-square test of the null hypothesis that the simple proportionality function and the power function provide equivalent descriptions of the relationship between the mean and the variance of cortical cell responses. The first bin (for monkey and cat) represents the number of cells which were better fitted by the power function, at the 0.05 level (approximately 20% averaged across all dimensions and all cells).

than would be expected by chance alone. However, even for those cells where the power function provided a significantly better fit, the visual goodness of fit was usually only marginally improved. Further, we have computed detection, discrimination, and identification performances for all the cells using both the proportionality function and the power function; generally, there was very little difference in the computed performances. We conclude that the simple proportionality function is the more parsimonious description of the relationship between the mean and the variance of cortical neuron responses.

As shown in the results (see Figs. 2–4), the value of K remains relatively constant within and across dimensions. The one exception was that for approximately half of the cells K tended to increase at lower temporal frequencies. Fig. 15 shows the value of K as a function of temporal frequency, averaged across all cells. The solid curve shows the best-fitting hyperbolic function:

$$K(w) = \frac{A}{w + B} + C \tag{B6}$$

where A, B, and C are free parameters and w is temporal frequency. As can be seen, K is relatively constant except at lower temporal frequencies. We have analyzed this relationship using trials consisting of a fixed amount of time, as well as a fixed number of cycles; the increase in K at low temporal frequencies was observed for both methods of analysis.

To quantify the magnitude of the variation in K, eqn. (B6) was fitted to each cell individually using the maximum-likelihood method, and then the value of K at 1.5 Hz was divided by the value of K at 15 Hz. The inset in Fig. 15 shows the distribution of this index. The dark bars indicate the cells for which the hyperbolic function provided a significantly better fit than a flat line, by a chi-square test.

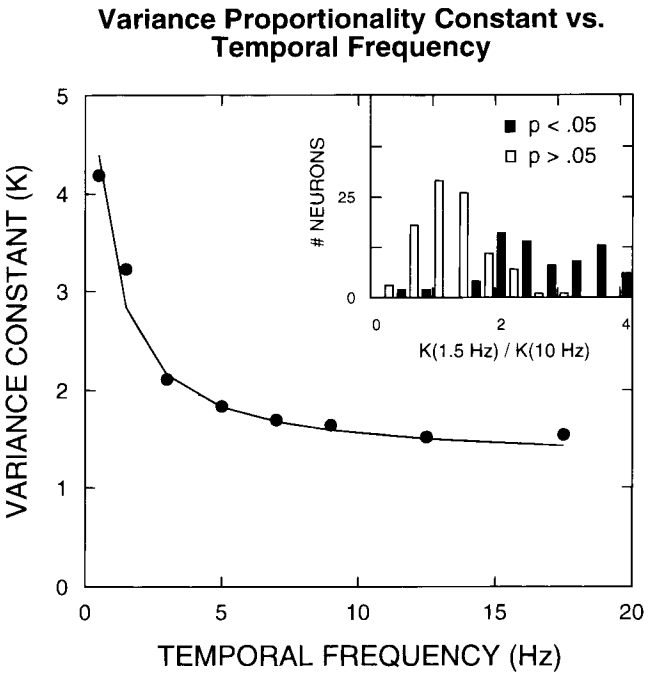

Fig. 15. Average K and temporal frequency. The variance proportionality constant, K, plotted as a function of temporal frequency, averaged across all of the cells. The smooth curve is the best-fitting hyperbolic function [see eqn. (B6)]. The histogram shows the magnitude of the effect for each cell considered individually: the value of K at 1.5 Hz divided by the value at 15 Hz. The black bars are the cells for which the variation in K was significant at the 0.05 level.

Appendix C: Descriptive function method versus standard method

The methods we used to measure detection, discrimination, and identification performance require much less data than standard methods. This efficiency derives from the fact that all of the computations are based upon descriptive functions for the mean and variance. All that is required to measure an entire discrimination function, plus an entire set of identification confidence regions, is an amount of data sufficient to estimate the few parameters of a descriptive function.

An obvious question is how well the descriptive function method compares with more standard methods. We addressed this question in two ways. The first was to examine the robustness of the method to violations of the assumption that the pulse-number distributions can be approximated as normal distributions. The second was to compare discrimination functions derived from the same raw spike trains using the descriptive function method and a standard method.

The pulse-number distributions of cortical cells are not always well approximated by normal distributions. Typically, at low spike rates the distributions are skewed to the right (unlike the normal distribution), whereas at high spike rates they become more symmetric (like the normal distribution). The Poisson distribution has this same property and hence it is a good comparison distribution for evaluating the error introduced by the assumption of normality.[9] We computed discrimination functions, certainty functions, and identification confidence regions using Poisson pulse-number distributions and normal distributions that had the same means and variances as the Poisson distributions. We also computed population performance using the d' summation formula [eqn. (3)]. The results showed that the functions obtained for the two distributions were very similar—the errors were almost always under 5%.

A more direct test of the adequacy of a new method is to compare it with an accepted standard method on the same set of spike data. Geisler et al. (1991) used standard methods to measure contrast discrimination functions for three cortical cells. In that analysis, each stimulus contrast was presented on 200 separate occasions to provide a reasonable estimate of the pulse-number distribution for each contrast. These measured distributions were then used to determine neurometric functions. The contrast discrimination thresholds (the 70% correct points on the neurometric functions) are shown as the solid symbols in Fig. 16. In comparison, the solid curves show the contrast discrimination functions obtained using the descriptive function method. For this latter analysis, the first 40 presentations of the same data set were used to estimate the parameters which describe the response as a function of contrast. As can be seen, there is good agreement between the two methods. The deviations seen at saturating contrasts are due to the fact that the location along the contrast axis where threshold increases steeply is very sensitive to the exact shape of the contrast response function in the region of saturation. Because the descriptive function method smoothly interpolates all the data in the saturation region it may be more accurate.

Appendix D: Contrast-gain exponent model

A model of cortical simple cells [the Contrast-Gain Exponent (CGE) model] was used to obtain responses to the complex image in Fig. 13. This model (see Fig. 12) is similar to those proposed by Heeger (1991) and Albrecht and Geisler (1991, 1994), and is consistent with the descriptive functions in Table 1 (i.e. the model predictions are described by the formulas in Table 1, for sine-wave stimuli). The equations and parameters defining the model are given here. For a related psychophysical model see Foley (1994).

For this model, the mean response of a cortical cell is given by the following equation:

[9] Although the Poisson distribution is asymmetric for small means, similar to cortical cells, it does not provide a sufficiently general model for cortical cell variability because the variance proportionality constant, K, can only be 1.0.

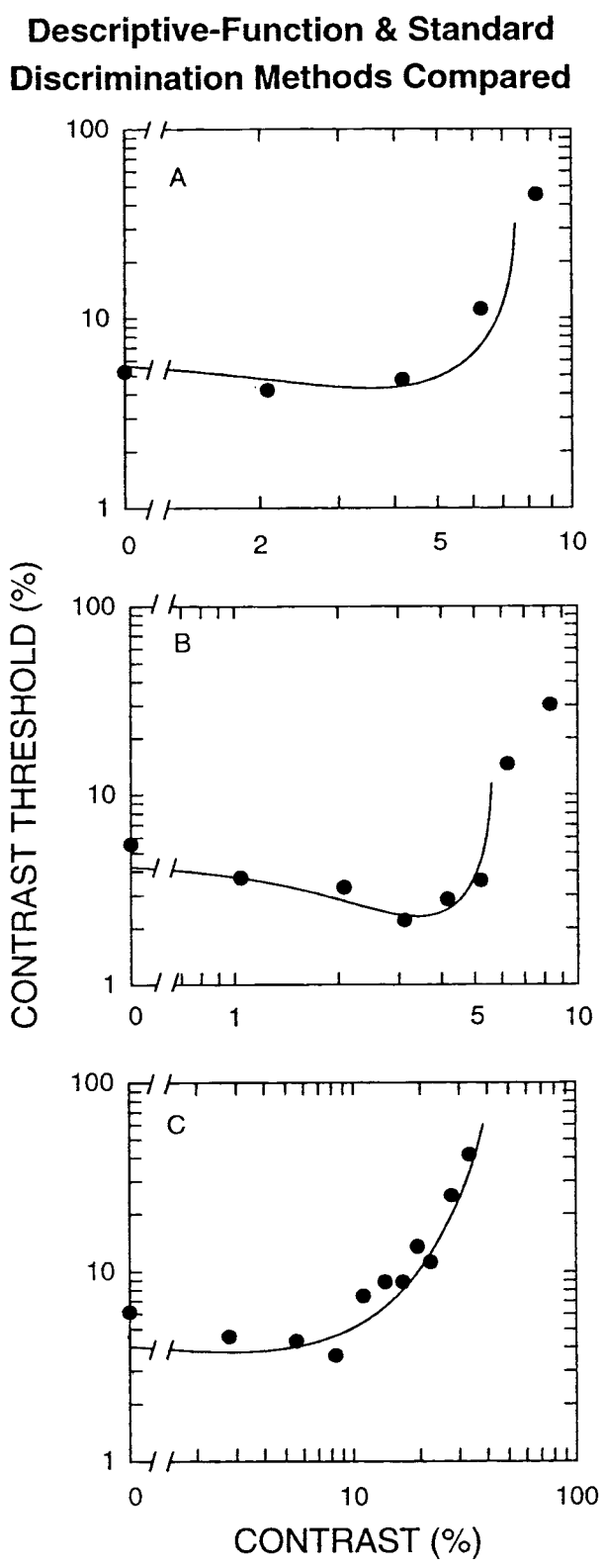

Fig. 16. Comparison of methods to measure discrimination performance. Comparison of two different methods for estimating contrast discrimination performance, using the same set of measurements on three separate cells (from Geisler et al., 1991). For the standard method (solid symbols), 200 repeated presentations at each contrast were utilized to estimate the pulse-number distributions and neurometric functions. For the descriptive function method (solid lines), only the first 40 repeated presentations at each contrast were utilized to estimate the mean and the standard deviation as a function of contrast. As can be seen, the two methods provide similar results.

$$r = r_{\max} \frac{\max\left(\sum_i c(x_i, y_i) h(x_i, y_i), 0\right)^n}{\bar{c}^n + c_{50}^n} + r_0 \qquad \text{(D1)}$$

where $c(x_i, y_i)$ is the effective contrast of the ith image point (e.g. pixel), $h(x, y)$ is a linear weighting function, $\bar{c}$ is the local mean contrast, $r_{\max}$ is the maximum response to the optimal sine-wave grating, c_{50} is the half-saturation contrast, n is the response exponent, and r_0 is the base rate. Note that applying the max function in the numerator, with the second argument set to zero, is equivalent to half-wave rectification.

The linear weighting function, $h(x, y)$, describes the linear receptive-field properties of the neuron; it determines the basic spatial frequency and orientation selectivity of the cell. (Temporal selectivity and direction selectivity can be incorporated into the model by including the dimension of time in the weighting function.)

The effective contrast is defined by the following equation:

$$c(x_i, y_i) = \frac{l(x_i, y_i) - \bar{l}}{\bar{l} + l_{50}} \qquad \text{(D2)}$$

where $l(x_i, y_i)$ is the luminance of the ith image point (e.g. pixel), $\bar{l}$ is the local mean luminance, and l_{50} is the light-adaptation constant. This definition of effective contrast incorporates multiplicative and subtractive light-adaptation into the model. These adaptation mechanisms are needed to account for the fact that cortical cells primarily respond to contrast rather than luminance. Examination of eqn. (D2) shows that the multiplicative adaptation factor is $1/(\bar{l} + l_{50})$; the subtractive adaptation factor is $\bar{l}/(\bar{l} + l_{50})$ (assuming subtractive adaptation occurs second).

The local mean luminance, $\bar{l}$, is defined by

$$\bar{l} \sum_i f(x_i, y_i) l(x_i, y_i) \qquad \text{(D3)}$$

where $f(x, y)$ is a nonnegative weighting function describing the spatial extent over which the local mean luminance is determined. Similarly, the local mean contrast, $\bar{c}$, is defined by

$$\bar{c} = \sqrt{\sum_i [g(x_i, y_i) c(x_i, y_i)]^2} \qquad \text{(D4)}$$

where $g(x, y)$ is a nonnegative weighting function describing the spatial extent over which the local mean contrast is determined.[10]

The variance of the response is given by

$$\sigma^2 = Kr \qquad \text{(D5)}$$

where K is the variance proportionality constant, and r is the mean response from eqn. (D1).

Eqns. (D1)–(D5) are the formulas used to generate the predictions of the model. Consider the relationship of these formulas to the diagram in Fig. 12. The curve within the "contrast gain box" of Fig. 12 is the contrast normalization function, which describes how the multiplicative contrast gain factor varies as a function of the local contrast. Inspection of eqn. (D1) shows that this function is of the form

$$\mu(\bar{c}) = \frac{1}{\sqrt[n]{\bar{c}^n + c_{50}^n}} \qquad \text{(D6)}$$

The curve within the "linear summation box" of Fig. 12 represents the Fourier transform of the linear receptive field, $h(x, y)$ (only one spatial dimension is represented). The curve within the "response exponent box"

[10] A more general version of the model would allow the contrast signal, $c(x, y)$, which feeds the contrast normalization mechanism, to be filtered prior to averaging. This would allow the model to incorporate spatial or temporal tuning of the normalization.

of Fig. 12 is the response exponent function, which describes how the final response varies with the input from the prior stage. This curve is a power function with an exponent of n:

$$v(x) = x^n \tag{D7}$$

Finally, the curve within the "multiplicative noise box" of Fig. 12 represents the relationship between the response variance and the response mean: it is a linear function with a slope of K and an intercept of 0.0.[11]

To generate the predictions in Fig. 13, we first constructed a large array of model cortical cells. The locations of the receptive-field centers fell on a regular grid with a spacing approximately equal to the spacing of on-center midget ganglion cells in the human eye (Curcio & Allen, 1990; Geisler & Banks, 1995). Each grid location contained 12 randomly generated cortical cells. For each model cortical cell, the half-saturation contrast, the response exponent, and the receptive-field orientation were randomly sampled from probability distributions based upon published frequency histograms for these parameters and the data summarized in Table 2 (Seay et al., 1996). The linear receptive-field weighting functions were sine-phase Gabor functions with a circularly symmetric envelope and bandwidth of 2 octaves. Because of the variation in the response exponent the effective bandwidths of the neurons ranged from less than 1 octave to greater than 2 octaves. We only generated predictions for the population of neurons with a peak spatial frequency of 16 cpd at the center of the fovea (i.e. for cells carrying detailed form information). Adding a wide range of peak spatial frequencies resulted in plots that were too cluttered to be easily interpreted. To represent retinal/cortical magnification effects, the peak spatial frequency varied inversely with the grid spacing (i.e. the spacing of the ganglion cells). The local weighting functions for luminance and contrast were taken to be the envelope of the linear receptive field, $e(x, y)$, normalized appropriately:

$$f(x, y) = \frac{e(x, y)}{\sum_i e(x_i, y_i)} \tag{D8}$$

$$g(x, y) = \sqrt{2}\,\frac{e(x, y)}{\sqrt{\sum_i e^2(x_i, y_i)}} \tag{D9}$$

(The $\sqrt{2}$ factor scales the local mean contrast so that it agrees with the Michelson contrast for sine-wave gratings.) The peak response to sine-wave gratings, r_{max}, was set to 10, and the base rate, r_0, to 0.0. The number of spikes produced by each cell in a 200-ms analysis interval was computed by randomly sampling from a normal distribution with a mean given by eqn. (D1), and a variance given by eqn. (D5) with $K = 1.4$.

Although some simplifying assumptions were made in generating the responses shown in Fig. 13, they should have no effect on the major conclusion: even though individual cortical neurons are broadly tuned and quite noisy, they do provide good identification information about local image features when they produce responses at or above 50% of r_{max} (i.e. 5–10 spikes).[12]

[11] One possible objection to this version of the CGE model is that the final response exponent appears in the contrast normalization function. However, there are essentially equivalent versions of the model that do not have this property. For example, we have found that a normalization function of the form:

$$\mu(\bar{c}) = \frac{1 - e^{-\alpha\bar{c}}}{\alpha\bar{c}}$$

gives nearly identical fits to single cell data (although the final response exponent is a little larger). We chose to use eqn. (D6) for historical reasons and to keep the CGE model consistent with the descriptive formulas in Table 1. The descriptive contrast response function used with the above normalization function is

$$r_c(c) = (1 - e^{-\alpha c})^n$$

[12] Results similar to those in Fig. 13 were obtained when we set all of the parameters, except orientation, to the mean values reported in the literature, rather than randomly sampling from distributions. One minor difference was that fewer low contrast contours were detected (because the average value of c_{50} was approximately 0.4). Note, however, that this version of the model does not account for other results in the report (e.g. Figs. 10 and 11).