

Revisiting spatial vision: toward a unifying model

Laurent Itti and Christof Koch

California Institute of Technology, Computation and Neural Systems, Mail Code 139-74, Pasadena, California 91125

Jochen Braun

Institute for Neuroscience, University of Plymouth, Plymouth, PL4 8AA, UK

Received February 24, 2000; revised manuscript received May 10, 2000; accepted July 6, 2000

We report contrast detection, contrast increment, contrast masking, orientation discrimination, and spatial frequency discrimination thresholds for spatially localized stimuli at 4° of eccentricity. Our stimulus geometry emphasizes interactions among overlapping visual filters and differs from that used in previous threshold measurements, which also admits interactions among distant filters. We quantitatively account for all measurements by simulating a small population of overlapping visual filters interacting through divisive inhibition. We depart from previous models of this kind in the parameters of divisive inhibition and in using a statistically efficient decision stage based on Fisher information. The success of this unified account suggests that, contrary to Bowne [Vision Res. **30**, 449 (1990)], spatial vision thresholds reflect a single level of processing, perhaps as early as primary visual cortex. © 2000 Optical Society of America [S0740-3232(00)02311-5]
OCIS codes: 330.0330, 330.1800, 330.4060, 330.5510, 330.6100, 330.7310.

1. INTRODUCTION

Early visual processing is commonly modeled with a population of visual filters selective for different locations, orientations, and spatial frequencies.¹⁻⁹ This approach not only is based on psychophysical evidence from a variety of paradigms, such as selective adaptation,¹⁰ contrast masking,^{6,11,12} and subthreshold summation,^{13,14} but also agrees with physiological evidence that neuronal responses in visual cortex are selective for different locations, orientations, and spatial frequencies.¹⁵⁻¹⁷ Indeed, the visual filter properties inferred from psychophysics agree reasonably well with the receptive field properties of neurons in primary visual cortex.¹⁸⁻²⁰

The first models of this kind assumed that visual filters respond independently of each other. Given independence and some assumption about noise, a wide range of visual detection and discrimination thresholds can be derived from the tuning properties of the filters. This includes thresholds for detecting or discriminating pattern contrast,^{3-6,11} thresholds for detecting a pattern in the presence of a superimposed mask,^{6,11,12,21} thresholds for discriminating pattern orientation or spatial frequency,²²⁻²⁴ and thresholds in a variety of hyperacuity tasks, among them Vernier acuity, chevron acuity, and line bisection acuity.²⁵⁻²⁷ In each case the just-discriminable difference between two stimuli can be derived from the just-detectable variation in filter responses.

More recent evidence shows, however, that visual filters are independent only to a very first approximation. Interactions between spatially adjacent but nonoverlapping visual filters (surround interactions) have been demonstrated both psychophysically²⁸⁻³⁰ and physiologically for neurons in primary visual cortex.³¹⁻³⁶ Surround interactions may be either excitatory or inhibitory and

are likely to contribute to texture segmentation^{37,38} and to visual grouping.³⁹⁻⁴¹

In addition, there is good evidence for interactions between overlapping visual filters (local interactions). The most relevant paradigm to this issue is contrast masking, where the contrast threshold of one pattern (target) is measured in the presence of another, superimposed pattern (mask).^{9,42-45} The fact that masks of different shapes are not interchangeable (even when mask contrast is suitably adjusted) constitutes strong evidence for interactions.^{9,42,43} In general, the effect of a mask declines as a measure of the difference in the orientation or spatial frequency relative to the target.^{6,11,12} In addition, the mask effect may be excitatory or inhibitory, depending on whether the mask has low or intermediate-to-high contrast (dipper function^{4,6,46}). Another consistent observation is that excitatory mask effects abate more rapidly than inhibitory ones when target and mask differ in phase or orientation, suggesting that excitatory interactions involve a narrower range of visual filters.^{9,43,47}

Physiological evidence for local interactions comes from studies of cross correlation in neuronal firing,^{48,49} contrast adaptation,⁵⁰⁻⁵⁴ and contrast masking^{55,56} in primary visual cortex. Interestingly, some of the psychophysical and physiological models of local interactions are strikingly similar,^{9,19,57-61} not least in that they formalize the interaction in terms of divisive inhibition.^{51,52}

The present paper aims to characterize local interactions, while avoiding the potentially confounding effects of surround interactions. To this end, we conducted contrast-masking experiments with spatially localized stimuli that are presented in the visual periphery. Given the cortical magnification of primary visual cortex at the relevant eccentricity, we expect that our stimuli will be processed by a single hypercolumn (a set of neu-

rons tuned to a range of orientations and spatial frequencies, with spatially overlapping receptive fields) in primary visual cortex.^{15,62} Most previous studies of contrast masking have used spatially extended stimuli (at least for the mask) and thus do not distinguish between local and surround interactions. As far as we know, only two previous studies used spatially localized stimuli, but they presented them in the fovea where cortical magnification is far higher.^{43,44} Thus the stimuli in question are still processed by dozens of hypercolumns and are subject to both local and surround interactions. By targeting a smaller patch of visual cortex, we hoped to generate more tractable results that can be modeled in terms of local interactions only.

In addition to contrast-increment and contrast-masking thresholds, we also established orientation and spatial frequency discrimination thresholds for identical stimuli. From a computational point of view, one would expect the two types of thresholds to reflect the same early level of visual processing.^{20,63,64} However, the quantitative consistency of discrimination and masking thresholds has never been established. Psychophysical studies typically report either orientation/spatial frequency discrimination thresholds^{22,23,65–67} or contrast-increment/masking thresholds,^{6,9,11,12,21,43–45,57,58} but not both. Indeed, the few studies that depart from this rule were unable to account simultaneously for both types of thresholds.^{27,68} The main difficulty is that the two types of thresholds exhibit different dependences on stimulus contrast.

We present a model that simultaneously accounts for contrast-increment/masking and orientation/spatial frequency discrimination thresholds measured in five separate experiments. This demonstrates that a wide range of spatial vision thresholds reflect the same early level of visual processing, presumably in or near primary visual cortex. Our model is similar to others in that it is based on divisive inhibition among a population of visual filters.^{8,9,43,45,52,53,57,58,69} The main difference lies in the parameters of divisive inhibition, in the use of a statistically efficient decision stage, and in the comprehensive analysis of the model's ten-dimensional parameter space. We have shown elsewhere that this model accounts also for the modulation of visual thresholds by attention.⁷⁰ By attempting to formulate a consensus between previous models for psychophysical and neuronal sensitivity to contrast, orientation, and spatial frequency, with our model we try to link visual thresholds more closely to the underlying mechanisms in visual cortex.

2. METHODS

A. Psychophysics

We studied the discrimination of Gabor stimuli in the near periphery with respect to contrast, orientation, and spatial frequency. Seven naïve observers participated in the experiments. For each task, observers received 15–24 blocks (of 100 trials each) of training and collected data for 45–90 blocks. Each threshold estimate reflects between 10 and 18 blocks. Observers carried out 10–12 blocks in a single, hour-long session, typically shared

between two tasks (Experiments 1 and 2, Experiments 1 and 3, or Experiments 4 and 5, in this order).

Stimuli were generated with an SGI Indigo workstation. Color-bit stealing⁷¹ was used to reduce the minimum luminance step of the display from 1.5% to 0.2%. Screen luminance varied from 1 to 90 cd/m² (mean 45 cd/m²), and room illumination was 5 cd/m². Displays subtended 16° × 13° of visual angle for a viewing distance of 80 cm. All tasks employed a temporal two-alternative forced-choice (2AFC) paradigm. Trials consisted of a circular cue (1° diameter, 250 ms), a 300-ms blank interval, and the two alternative targets in random order (250 ms each), separated by a second 300-ms blank interval (total duration 1350 ms). Both targets appeared at the cued location. Observers reported the order of alternative targets by pressing one of two buttons and received auditory feedback. During each block of 100 trials, the alternative targets were modified following a staircase procedure (reducing target similarity after two consecutive mistakes and increasing target similarity after four consecutive successes). Threshold was taken to be the 75% correct performance level of a Weibull function $\{1 - 1/2 \exp[-(\alpha x)^\beta]\}$, with $\beta = 2$ and α determined by a maximum-likelihood fit to the psychometric data.^{72,73} Four percent of all blocks were excluded because the staircase failed to converge.

Targets appeared at a constant eccentricity of 4° but at random polar angles. To further discourage eye movements, observers were instructed to fixate a cross at display center. The targets were Gabor stimuli or superpositions of Gabor stimuli [spatial frequency 1.4 to 5.6 cycles per degree (c/deg); half-width at half-maximum equal to the spatial period]. At 4° eccentricity, cortical magnification is approximately 1.7 mm/°, compared with approximately 10.0 mm/° in the fovea.⁷⁴ Thus the stimulus diameter of approximately 0.7° (2 cycles at 2.8 c/deg) corresponds to approximately 1.2 mm in primary visual cortex, that is, less than the average diameter of a hypercolumn.⁷⁵ Presented in the fovea, the same stimulus would have excited approximately 35 hypercolumns.

Experiment 1 (Exp. 1) involved Gabor stimuli of different contrast (2.8 c/deg, vertical orientation, cosine phase). Observers reported which target had higher contrast, and threshold was established in terms of incremental contrast.^{4,6,9} Experiment 2 (Exp. 2) concerned Gabor stimuli of different orientation (2.8 c/deg, contrast 0.05–0.9, cosine phase). Observers reported whether the target was tilted clockwise or counterclockwise, and threshold was measured in terms of tilt angle.⁷⁶ Experiment 3 (Exp. 3) used targets of different spatial frequency (vertical orientation, contrast 0.05–0.9, cosine phase). One target was at 2.8 c/deg, and the other target exhibited lower (coarser) spatial frequency. Observers reported which target had lower spatial frequency, and threshold was measured in terms of the spatial frequency ratio in octaves.⁷⁶ Two further experiments concerned the discrimination of two superimposed Gabor stimuli. One was present in both intervals (“mask,” contrast 0.5, random phase, uniformly distributed between 0° and 360°), and the other appeared in only one interval (“target,” 2.8 c/deg, vertical orientation, cosine phase). The random phase of the mask ensured that observers could not rely

on any specific feature formed by the superposition of two stimuli. Observers reported which interval contained the target stimulus, and threshold was measured in terms of target contrast. In Experiment 4 (Exp. 4), mask orientation was varied between blocks (0° – 90°), but spatial frequency remained fixed (2.8 c/deg).¹² In Experiment 5 (Exp. 5) mask spatial frequency was varied between blocks [0.5–2 octaves (oct.)], but orientation was fixed (15° from vertical).¹¹

B. Modeling

The model comprises three successive levels: linear filtering, divisive inhibition, and a statistically efficient decision strategy (Fig. 1).

1. Linear Filters

We use a population of overlapping spatial filters, all centered at the same point of visual space but tuned to a variety of spatial periods $\lambda \in \Lambda$ and orientations $\theta \in \Theta$. To facilitate comparison with cortical neurons and increase computational efficiency, filters are defined in terms of their tuning functions rather than in terms of the spatial structure of their receptive fields. For sinusoidal grating stimuli, we assume Gaussian tuning with respect to both the logarithm of stimulus spatial period λ_S (with standard deviation σ_λ) and orientation θ_S (with standard deviation σ_θ). Thus the response of a filter with preferred period λ and preferred orientation θ to a sinusoidal grating of contrast C_S , period λ_S , and orientation θ_S is given by

$$E_{\lambda,\theta}(C_S, \lambda_S, \theta_S) = C_S A \exp\left\{-\frac{[\log(\lambda_S) - \log(\lambda)]^2}{2\sigma_\lambda^2} - \frac{(\theta_S - \theta)^2}{2\sigma_\theta^2}\right\}, \quad (1)$$

where A is a gain coefficient. Note that this definition specifies filters in the Fourier domain and disregards phase information. To obtain responses to arbitrary stimuli, we realized pairs of these filters with quadrature phase^{56,77} and combined their outputs by taking the square root of the sum of the squared responses. The reconstructed spatial filters closely resemble Gabor func-

tions and neuronal receptive fields and are illustrated in Fig. 2. Note that for the Gabor stimuli used here, filter responses can also be computed directly from the tuning functions.

2. Divisive Inhibition

The response of each linear filter is normalized relative to the total population response (divisive inhibition⁵²).^{8,9,43,45,78} The functional consequences of divisive inhibition include a nonlinear transducer^{4,46} and reduced contrast dependence of orientation and spatial frequency tuning.⁸ The normalized response $R_{\lambda,\theta}$ of a filter tuned to (λ, θ) is

$$R_{\lambda,\theta} = \frac{(E_{\lambda,\theta})^\gamma}{(S)^\delta + \sum_{(\lambda',\theta') \in \Lambda \times \Theta} W_{\lambda,\theta}(\lambda', \theta') (E_{\lambda',\theta'})^\delta} + \eta, \quad (2)$$

where

$$W_{\lambda,\theta}(\lambda', \theta') = \exp\left\{-\frac{[\log(\lambda') - \log(\lambda)]^2}{2\Lambda_\lambda^2} - \frac{(\theta' - \theta)^2}{2\Lambda_\theta^2}\right\} \quad (3)$$

is a two-dimensional Gaussian weighting function centered around (λ, θ) whose widths are determined by the scalars Λ_θ and Λ_λ . In Eq. (2), η is a positive constant representing background activity (dark current). This constant is needed to avoid singularities in the decision stage for stimuli of zero contrast. The denominator includes a constant S and the weighted sum of all filter responses and represents divisive inhibition. The exponents γ and δ determine the resulting transducer function. Depending on their values, one obtains a sigmoidal ($S > 0$, $\gamma > \delta$), power-law ($S = 0$), or even linear ($\gamma = 1$, $\delta = 0$, $S = 0$) transducer, with a saturating ($\gamma = \delta$) or nonsaturating ($\gamma \neq \delta$) response.

It is convenient to replace A and S with two alternative parameters that are easier to interpret. We use the detection threshold C_{th} for a grating stimulus of 2.8 c/deg and the position of the transducer's inflexion point

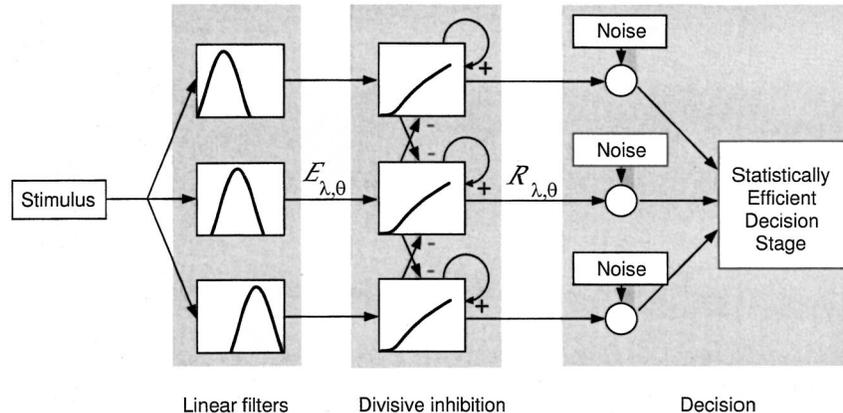


Fig. 1. Model architecture, represented schematically in the style of Wilson and collaborators. The model consists of three successive stages: (1) A bank of linear visual filters tuned to different orientations and spatial frequencies, (2) nonlinear interactions between visual filters in the form of a power law and divisive inhibition, and (3) addition of independent noise and a statistically efficient decision based on the entire filter population.

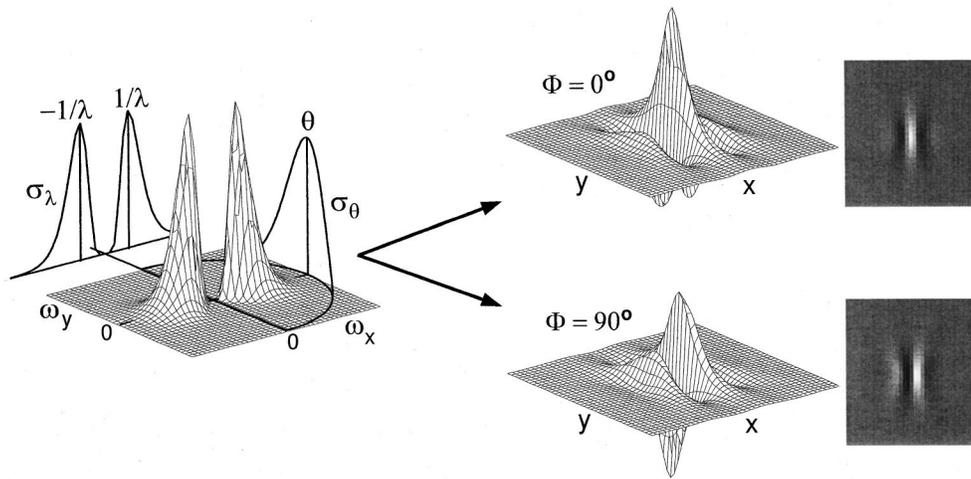


Fig. 2. Visual filters used in the model. Filters are defined in terms of their separable Gaussian tuning functions for the orientation θ and logarithm of the spatial period λ of a sinusoidal grating stimulus (left). The spatial shape of the visual filter can be reconstructed through an inverse Fourier transform. Both even- and odd-symmetric filters are shown (middle and right). Their shape is very similar to the multilobed functions used by other models. Pairs of reconstructed even- and odd-symmetric filters can be used to numerically compute the response to arbitrary stimuli. However, the response to stimuli based on sinusoidal gratings (e.g., Gabor patches) can be obtained directly from the Gaussian tuning functions.

$S_{\text{fac}}(S_{\text{fac}} = C_{\text{infect}}/C_{\text{th}})$. If all other model parameters are fixed, each choice of A, S corresponds to a unique choice of $C_{\text{th}}, S_{\text{fac}}$, through a closed-form expression not detailed here.

3. Noise Model: $Poisson^\alpha$

Following normalization, independent Gaussian noise is added to each filter response. In analogy to visual cortical neurons, we assume that the variance increases with the response mean:

$$V_{\lambda, \theta}^2 = R_{\lambda, \theta}^\alpha, \tag{4}$$

where α is a constant. For visual cortical neurons, α is typically slightly larger than unity.^{20,79,80} Note that this noise assumption differs from most psychophysical models, which assume Gaussian noise of constant variance.^{3,6,7,45,81}

4. Decision Stage

We use a statistically efficient decision stage to predict behavioral thresholds from the noisy responses of model units. This ideal-observer decision accounts significantly better for our psychophysical data than the nonideal decision strategies used in other models (see discussion of results in Section 3).

We consider the noisy response vector $\mathcal{R} = \{R_{\lambda, \theta}; \lambda \in \Lambda, \theta \in \Theta\}$, with a mean given by Eq. (2) and a variance by Eq. (4). For each stimulus attribute ζ , we postulate a statistic $T(\zeta)$ that estimates the value of ζ from the noisy response \mathcal{R} . The discrimination performance in a 2AFC experiment involving two stimuli with attributes ζ_1 and ζ_2 is given by⁸¹

Prop. Corr. =

$$\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{|\operatorname{mean}[T(\zeta_1)] - \operatorname{mean}[T(\zeta_2)]|}{\sqrt{2\{\operatorname{var}[T(\zeta_1)] + \operatorname{var}[T(\zeta_2)]\}}} \right), \tag{5}$$

where erf is the normal error function. We assume that $T(\zeta)$ is “ideal” in the sense that it is both unbiased and

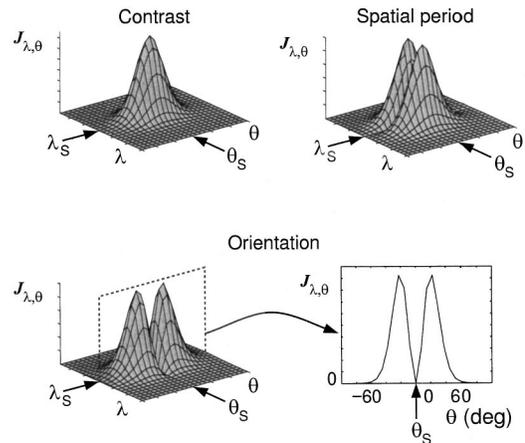


Fig. 3. Fisher information with respect to stimulus orientation, contrast, and spatial frequency. Fisher information is the inverse of the variance of an unbiased and efficient estimator of the stimulus parameter [Subsection 2.B.4. and Eq. (7)]. Each surface point represents the information encoded in the response of one model unit. The volume under the surface represents the total information encoded by a population of units with independent noise and tuned to 24 orientations and 24 spatial periods. Arrows indicate the spatial period λ_S and orientation $\theta_S = 0$ of the stimulus. Note that the unit tuned optimally for the stimulus does not contribute to the Fisher information for orientation or spatial period.

efficient. The lack of bias implies that the estimator T exhibits no systematic bias toward either higher or lower values of ζ ; in other words, $\operatorname{mean}[T(\zeta)] = \zeta$. Efficiency implies that $\operatorname{var}[T(\zeta)]$ reaches the Cramér–Rao bound, the theoretical lower bound for the variance of any unbiased estimator.^{82,83} The Cramér–Rao bound equals the inverse of the Fisher information $J(\zeta)$, such that $\operatorname{var}[T(\zeta)] = 1/J(\zeta)$. It follows that

$$\operatorname{Prop. Corr.} = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{|\zeta_1 - \zeta_2|}{\sqrt{2[1/J(\zeta_1) + 1/J(\zeta_2)]}} \right). \tag{6}$$

The Fisher information $J_{\lambda,\theta}$ for a single unit with response $R_{\lambda,\theta}$ and variance $V_{\lambda,\theta}^2 = R_{\lambda,\theta}^\alpha$ is (Appendix A)

$$J_{\lambda,\theta}(\zeta) = \left(\frac{\partial R_{\lambda,\theta}}{\partial \zeta} \right)^2 \left(\frac{1}{R_{\lambda,\theta}^\alpha} + \frac{\alpha^2}{2R_{\lambda,\theta}^{2\alpha}} \right). \quad (7)$$

In other words, information is distributed over the population, and the units that respond maximally do not necessarily provide the most information. As the Fisher information is additive in the case of independent noise,^{83,84} the total Fisher information for the entire population of units simply is

$$J(\zeta) = \sum_{\lambda,\theta} J_{\lambda,\theta}(\zeta). \quad (8)$$

Note that this approach does not require a rule governing the decision on each particular trial. Effectively, the rule involves forming the difference $s_1 - s_2$ between two random variables s_1, s_2 and comparing it with zero, where $\text{mean}[s_{1,2}] = \text{mean}[T(\zeta_{1,2})]$ and $\text{var}[s_{1,2}] = \text{var}[T(\zeta_{1,2})]$.

Figure 3 illustrates the distribution of Fisher information computed for stimulus contrast, orientation, and spatial frequency across units with different tuning properties. Note that information about each stimulus attribute is concentrated in different subpopulations of units. Owing to our relatively sparse distribution of model units, we compute thresholds numerically (through iterative adjustments of ζ until threshold performance is reached). The closed-form solutions that hold for dense distributions of model units (Appendix B and Refs. 85 and 86) are less accurate, especially near contrast threshold.

This decision stage can readily be generalized to other psychophysical paradigms. For example, performance in a yes/no paradigm can be obtained by altering Eq. (6). Another possible generalization concerns decision uncertainty.^{87–90} In the present experiments there is no decision uncertainty, since always the same stimulus parameter (known to the observer) varies from trial to trial within each block. When this is not the case, and the decision involves several stimulus parameters, Eq. (4) has to be generalized to $\text{var}[T] = J(Z)^{-1}$, where Z is the vector of all relevant stimulus parameters.

5. Alternative Decision Stage

For comparison, we also used an alternative decision stage based on the Minkowski norm (“Quick probability summation,”^{9,12,24,68,91}). The discriminability of two stimuli with ζ_1 and ζ_2 is computed as

$$\mathcal{D}(\zeta_1, \zeta_2) = \left[\sum_{(\lambda,\theta) \in \Lambda \times \Theta} \left| \frac{R_{\lambda,\theta}(\zeta_2)}{V_{\lambda,\theta}(\zeta_2)} - \frac{R_{\lambda,\theta}(\zeta_1)}{V_{\lambda,\theta}(\zeta_1)} \right|^Q \right]^{1/Q}, \quad (9)$$

where Q is the Minkowski exponent (values above 3 yield similar results⁶⁸). Threshold is reached when $\mathcal{D} = 1$.

3. RESULTS

A. Psychophysics

Psychophysical results from seven observers are shown in Fig. 4. On average, the detection threshold for Gabor patches of 2.8 c/deg, subtending approximately 0.72° at 4° eccentricity, is reached at a contrast of approximately 0.025 (Fig. 4, Exp. 1). This is higher than the previously reported detection threshold of contrast 0.005 for gratings

of 2.0 c/deg subtending 0.75° and presented in the fovea.⁶ Presumably the discrepancy is due to the greater eccentricity^{92,93} of our stimuli.

The region of facilitation (dipper) occurs at a pedestal contrast of $S_{\text{fac}} \approx 1.2$ times the detection threshold C_{th} , and the lowest thresholds are two to three times smaller than C_{th} (Fig. 4, Exp. 1). This is in good agreement with previous studies, where the corresponding values are near 1.3 and 2.5, respectively.⁶ At higher pedestal contrast, increment thresholds increase with an exponent of approximately 0.6, consistent with the range of exponents (0.6 to 0.8) reported by previous studies.^{6,9,46} This close agreement is reassuring because it suggests that our peripheral thresholds reflect neural mechanisms similar to the foveal thresholds measured by previous authors.

In the limit of high contrast, average thresholds for orientation discrimination are approximately 2° (Fig. 4, Exp. 2) and average thresholds for spatial period discrimination approximately 0.05 oct. (Fig. 4, Exp. 3). Both values are at the upper end of the range reported previously for orientation discrimination (0.3°–2°) and spatial period discrimination (0.02–0.05 oct.) with relatively small patterns.^{22,24,65,94} Presumably the relatively high discrimination thresholds are due to the eccentric stimulus location.⁹²

Our experiments on contrast masking are similar to those reported by Wilson and colleagues.^{11,12} In our case the presence of a masking pattern of different orientation elevates contrast thresholds up to approximately fourfold (Fig. 4, Exp. 4), and a masking pattern of different spatial period elevates thresholds up to approximately 2.5-fold (Fig. 4, Exp. 5). The corresponding threshold elevations reported previously for stimuli presented in the fovea are approximately eightfold and fivefold, respectively.^{11,12} The difference may reflect either the smaller stimulus size or the greater stimulus eccentricity in our situation.

No threshold elevation is observed if target and masking patterns differ by more than 60° in orientation (Fig. 4, Exp. 4). This contrasts with previous reports of an approximately 2.5-fold threshold elevation with orthogonal target and masking patterns (cross-orientation inhibition).^{9,52} However, the latter studies used large masker stimuli (extending over 7° × 5° of visual angle in the fovea)⁹ and thus are likely to reflect interactions between both overlapping and nonoverlapping visual filters. Our stimuli were designed primarily to probe interactions between overlapping visual filters.

Three of our experiments measured contrast-increment threshold for patterns of 2.8 c/deg and a pedestal contrast of 0.5. In Exp. 1, target and mask patterns of identical orientation and phase produced an approximately 3.6-fold threshold elevation (Fig. 4, Exp. 1). In Exp. 4, target and mask patterns varied in relative phase, but in spite of this difference produced a comparable (approximately fourfold) threshold elevation (Fig. 4, Exp. 4). In Exp. 5, target and mask patterns differed in orientation (by 15°), resulting in a noticeably smaller (approximately 2.5-fold) threshold elevation (Fig. 4, Exp. 5).

B. Modeling

We implemented the model described above with 60 filters (spatial frequencies 1.4, 2, 2.8, 4, and 5.6 c/deg and

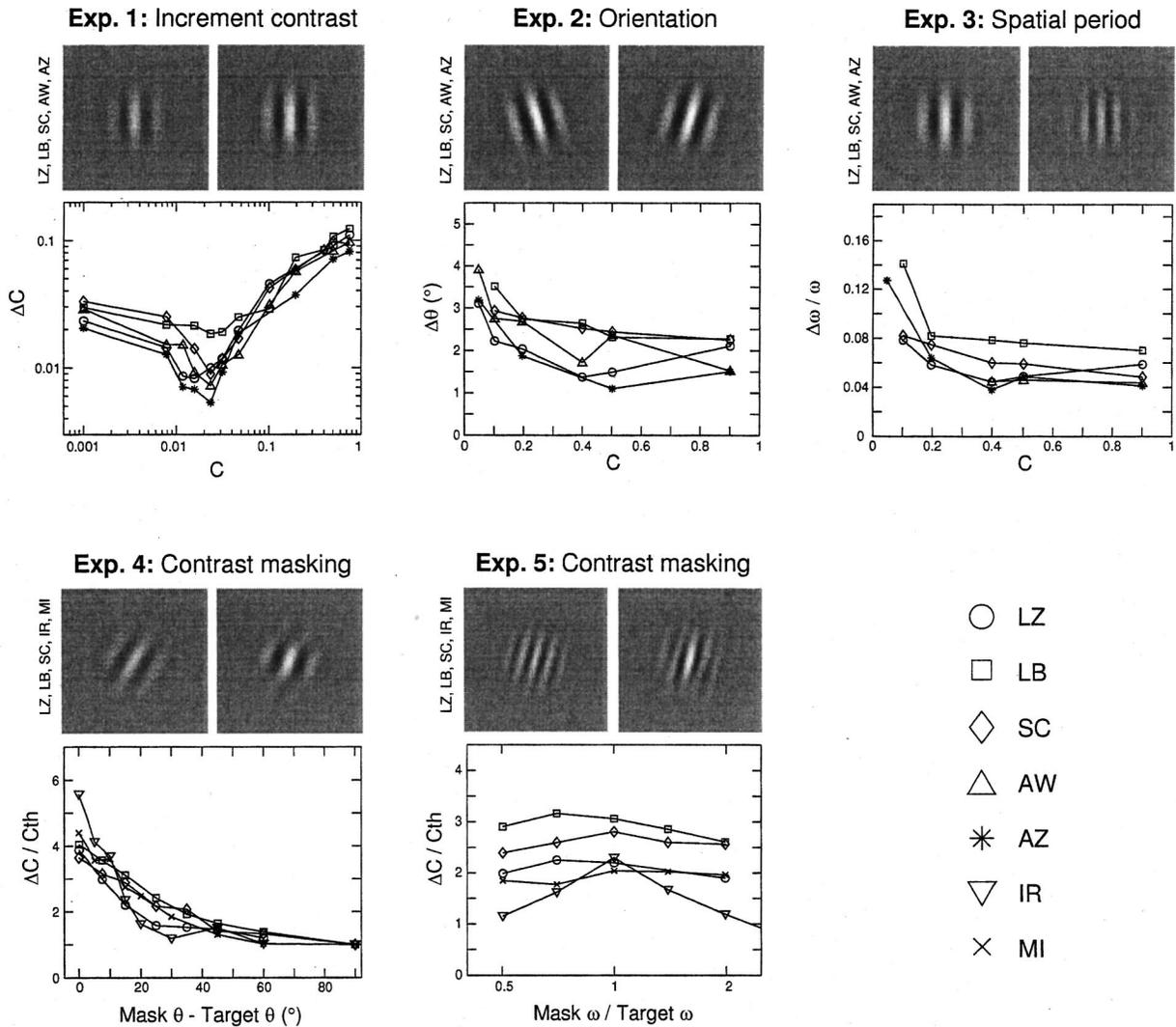


Fig. 4. Experimental results of seven psychophysical observers. Three observers completed all five experiments (LB, LZ, and SC), and others completed either two or three experiments (AW, AZ, IR, and MI). All experiments involve a temporal 2AFC discrimination between Gabor stimuli at 4° of eccentricity (insets). Exp 1: contrast increment threshold, ΔC , as a function of contrast, C . Exps. 2 and 3: orientation and relative spatial frequency discrimination threshold, $\Delta\theta$ and $\Delta\omega/\omega$, as a function of contrast, C . Exp. 4: contrast threshold elevation, $\Delta C/C_{th}$, as a function of mask orientation, θ (C_{th} is the detection contrast threshold, leftmost point of Exp. 1). Exp. 5: contrast threshold elevation, $\Delta C/C_{th}$, as a function of mask spatial period, λ .

orientations 0, 15, 30, ..., 165°). The parameters for gain A , background activity η , orientation tuning width σ_θ , and spatial frequency tuning width σ_λ (in octaves) were identical for all filters, resulting in an overall total of only ten free parameters (Table 1). Best-fitting parameter values were computed separately for each of the three subjects who completed all experiments (LZ, LB, and SC). The total fit error was computed as the root of the squared percentage errors summed over all data points. Using percentage errors rather than absolute errors ensured that all data points carried equal weight (i.e., independent of the unit of measurement).

Best fits were computed with two automatic procedures, with no operator bias toward “plausible” values. Both procedures produced consistent values. The first procedure used a ten-dimensional simplex algorithm⁹⁵ with simulated annealing overhead.^{96,97} From several randomly chosen starting points, the annealing schedule was initiated with a temperature, t , that induced random

parameter variations of $\pm 15\%$ at each transformation of the simplex. Every 200 simplex transformations, the annealing amplitude was reduced on a schedule $t \propto \log(1 + k)$, where k is the number of simplex iterations. The process was terminated when the annealing amplitude became smaller than $\pm 0.25\%$, at which point a final deterministic fit was carried out. The annealing schedule was sufficiently slow to ensure eventual convergence toward the global minimum.⁹⁸ The second procedure also used several randomly chosen starting points but approached the best fit with the help of Powell’s deterministic algorithm using Brent’s minimization method⁹⁵.

For each of the three complete data sets, the best fit of the model is shown in Fig. 5. The model accounts quantitatively for all observations, and in almost all cases predicted and measured thresholds agree to within the accuracy of the measurement. The model reproduced the dipper-shaped results of Exp. 1, the almost flat contrast dependence in Exps. 2 and 3, and the masking effects of

Exps. 4 and 5. The effective transducer function of the response was sigmoidal at low contrast and nonsaturating at high contrast [Fig. 6(a)], explaining the results of Exp. 1 and their conformance to Guilford's law,⁹⁹ which states that $\Delta C \propto C^x$, $0.5 \leq x \leq 1$. The effective tuning of the response for orientation was approximately 30% sharper than that of the linear units before the interactions (Fig. 2), and the effective tuning for spatial period was approximately 35% narrower [Figs. 6(c) and 6(d)], explaining the low thresholds for discriminating orientation and spatial period. The flat contrast dependence of these thresholds will be considered in a separate section.

The parameter values that yielded the best fit for each of the three complete data sets are listed in Table 1. Each value is given with a tolerance range (expressed as a percentage). This is the range in which a given parameter can vary such that the total fit error remains within 5% of its minimal value when all other parameters are optimized to keep the fit error as small as possible (Appendix C). Most tightly constrained (to within 2% of their respective values) were the exponents γ and δ . In the observer average, the optimal values were approximately $\gamma = 3.5$ and $\delta = 3.0$, that is, substantially larger than the values inferred from physiology.^{52,53} The tuning widths σ_θ , σ_λ , noise exponent α , contrast threshold C_{th} , and transducer inflection S_{fac} were constrained to within 15% of their respective values. The width of orientation tuning was approximately 35° full-width at half-maximum (FWHM) at the linear stage and approximately 25° at the nonlinear stage. Similarly, the width of spatial period tuning was approximately 1.5 oct. FWHM and 0.8 oct. FWHM at the linear and nonlinear stages, respectively. The noise exponent α was slightly larger than unity (values near 1.1). Less well constrained were the values for orientation pooling width Λ_θ (25° FWHM to within 4%–34%) and for background activity η (to within 18%–

185%). The spatial period pooling width Λ_λ was only weakly constrained by our data, mostly because the measured contrast-masking thresholds depended only weakly on the spatial period of the mask.

As the computation of tolerance values is based on the entire data set, it does not reveal the relative importance of different parts of the data set. To obtain some information on this point, we computed threshold predictions for the family of all models yielding fit errors less than 5% above the optimal fit (Appendix C). The envelope of the resulting predictions is narrow in the more important parts and broad in the less important parts of the data set (see Fig. 5). By this criterion, the most important parts of the data set are the high-contrast regimes of Exps. 1, 2, and 3 ($C > 0.1$). Experiments 4 and 5 appeared important as well, as the predicted envelopes are quite narrow. In the low-contrast regime of Exp. 1, the predicted envelopes are wide but concern only 3% of the entire contrast range of that experiment. Least important are the low-contrast regimes of Exps. 2 and 3, where predictions diverge widely. This analysis shows that the data did closely constrain the model and that, overall, the model was not sensitive to small departures from the optimal parameter values. This demonstrates that the optimal fit is robust and nonaccidental.

Finally, we were interested in knowing which parameter combinations are most critical for model predictions. To this end, we computed the eigenvectors of the Hessian of the error surface in parameter space and sorted them by their associated eigenvalues (Appendix C). For all three observers the largest eigenvalue was associated with a vector almost collinear with the contrast threshold C_{th} . This reflects the fact that C_{th} determines the overall sensitivity of the model and that its value modulates all predictions (both low- and high-contrast). The next largest eigenvalue was associated with the difference be-

Table 1. Best-Fit Model Parameters

Name	Symbol	Observer			SC	
		LB	LZ	SC	Constant Noise	Minkowski Norm
Excitatory exponent	γ	$3.2 \pm 5\%$	$3.8 \pm 1\%$	$3.6 \pm 1\%$	$3.5 \pm 1\%$	$4.9 \pm 5\%$
Inhibitory exponent	δ	$2.7 \pm 4\%$	$3.0 \pm 1\%$	$3.0 \pm 2\%$	$3.0 \pm 1\%$	$4.0 \pm 4\%$
Noise exponent	α	$1.1 \pm 40\%$	$1.3 \pm 2\%$	$1.0 \pm 25\%$	0	$0.8 \pm 7\%$
Background activity	η	$3.8 \pm 97\%$	$1.7 \pm 18\%$	$12.3 \pm 185\%$	$15.6 \pm 5000\%$	$1.7 \pm 262\%$
Spatial period tuning width	σ_λ (oct.)	$0.76 \pm 14\%$	$0.54 \pm 8\%$	$0.64 \pm 8\%$	$0.62 \pm 14\%$	$0.43 \pm 33\%$
Orientation tuning width	σ_θ (°)	$17.8 \pm 16\%$	$12.5 \pm 9\%$	$16.5 \pm 13\%$	$17.1 \pm 17\%$	$12.5 \pm 5\%$
Spatial period pooling width	Λ_λ (oct.)	$7.1 \pm 2499\%$	$5.7 \pm 602\%$	$1.2 \pm 661\%$	$0.6 \pm 933\%$	$1.4 \pm 4855\%$
Orientation pooling width	Λ_θ (°)	$14.5 \pm 5\%$	$12.0 \pm 34\%$	$11.0 \pm 4\%$	$10.8 \pm 38\%$	$9.3 \pm 7\%$
Contrast detection threshold	C_{th}	$0.026 \pm 1\%$	$0.025 \pm 3\%$	$0.026 \pm 1\%$	$0.025 \pm 1\%$	$0.023 \pm 1\%$
Transducer inflection point	S_{fac}	$1.26 \pm 1\%$	$0.62 \pm 8\%$	$0.80 \pm 8\%$	$0.80 \pm 1\%$	$1.37 \pm 23\%$
Residual fit error		16%	14%	17%	17%	20%
Linear spatial period FWHM ^a	(oct.)	1.79	1.27	1.51	1.46	1.01
Linear orientation FWHM	(°)	41.9	29.4	39.1	40.3	29.4
Pooled spatial period FWHM	(oct.)	1.12	0.68	0.71	0.86	0.43
Pooled orientation FWHM	(°)	27.5	18.0	26.1	28.0	18.5

^a Full width at half-maximum (FWHM) is computed as $2\sigma\sqrt{2\log(2)}$ for the linear filters and is measured at the output of the second stage (for a grating of contrast 0.1) for the pooled filters.

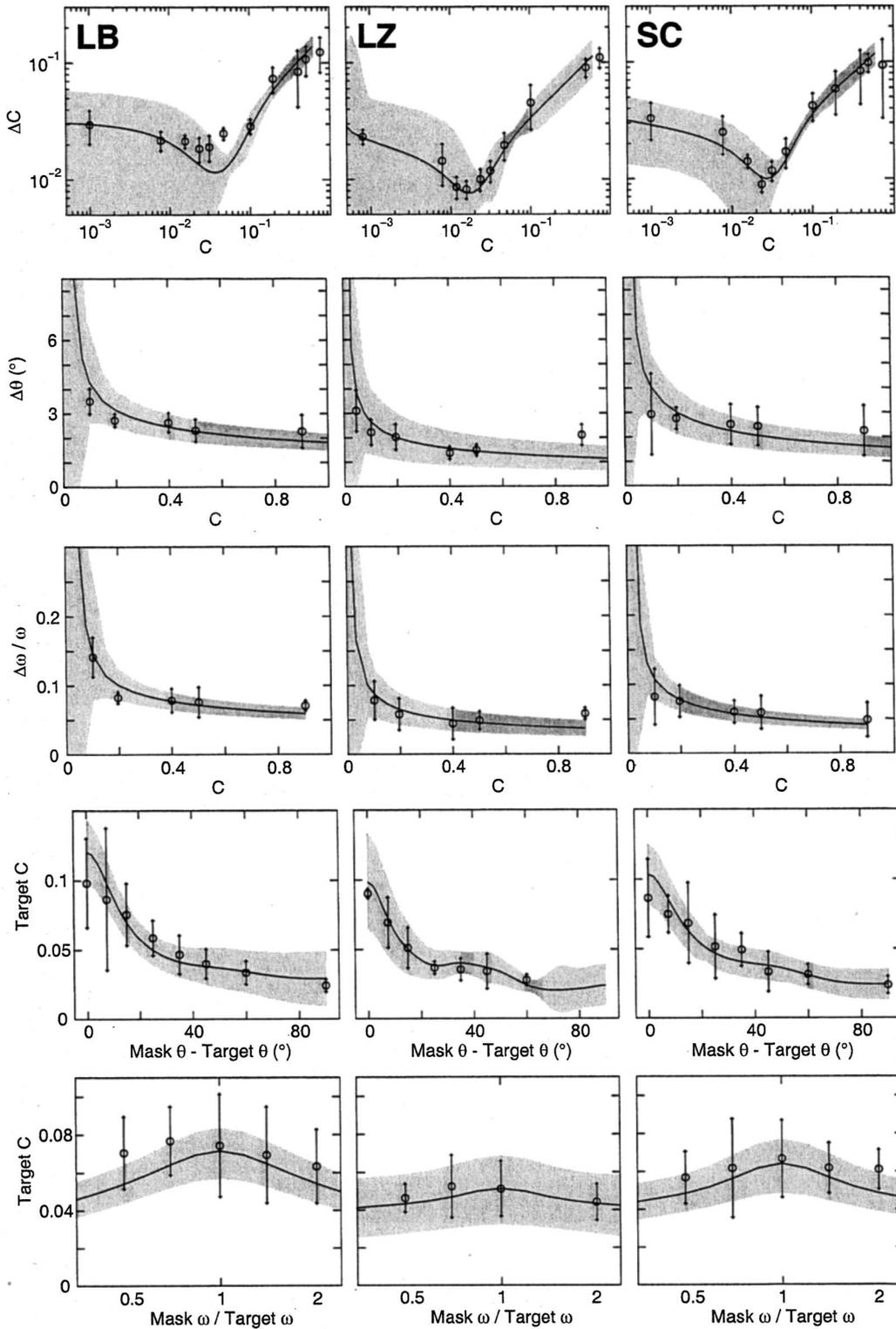


Fig. 5. Measured and predicted thresholds for three observers (LB, LZ, and SC). Measured thresholds (symbols with error bars) are represented by the mean and standard deviation for each observer. Predicted thresholds are represented by the optimal model fit (solid curve) and the family of all model fits with up to 5% higher fit error (gray regions). Predicted thresholds are close to measured thresholds, often to within the accuracy of the measurement. In general, the gray regions closely hug the measured thresholds, demonstrating that the model fit is robust and not accidental. The narrow parts of the gray regions indicate which parts of the data constitute particularly tight constraints for the model.

tween exponents, $\gamma - \delta$, which affects all high-contrast predictions and determines the asymptotic slope of the contrast response function in Fig. 6(a). Intermediate ei-

genvalues were associated with more complicated parameter combinations and were not consistent across subjects. However, in all observers the two smallest

eigenvalues were associated with the pooling widths Λ_θ and Λ_λ .

C. Variants of the Model

How does the number of model units (60 units with 12 preferred orientations and 5 preferred spatial frequencies) affect our conclusions? Given our decision model, which combines information from all units, the number of model units should be of little consequence as long as it is not too small. The reason is that the number determines how densely the Fisher information surface is sampled, but it does not alter its shape (see Fig. 3). Indeed, when we increased the number of units to 270 (30 orientations and 9 spatial frequencies), we observed no significant threshold changes except for an approximately 10% reduction in masking thresholds (Exps. 4 and 5). A further increase to 1296 units (72 orientations and 18 spatial periods) produced no significant change in the results (i.e., no prediction changes by more than 5%).

We depart from previous models by assuming that response variance increases with the response mean (Poisson noise, $V_{\lambda,\theta}^2 = R_{\lambda,\theta}^\alpha$ with $\alpha \approx 1$). A simpler alternative is to assume that response variance is constant (constant noise, $V_{\lambda,\theta}^2 = \beta$).^{9,24} Constant noise is a reasonable assumption as long as filters are independent and the contrast response follows a simple power law. In this case the dependence of the signal-to-noise ratio $R_{\lambda,\theta}/V_{\lambda,\theta}$ on stimulus contrast, orientation, and spatial frequency turns out to be the same for constant noise and Poisson noise (Appendix B). In the present case, however, filters interact through divisive inhibition, and the contrast response follows a sigmoidal law. Thus, in principle, Poisson noise and constant noise are no longer interchangeable. However, when we fit the data of observer SC with a variant of the model using constant noise, we obtain a good fit as well (Fig. 7). Nevertheless, the relatively

broad envelopes around the best fit indicate that the constant-noise model is less constrained by data and, in particular, that background activity (η) remains entirely unconstrained (Table 1).

Another difference from earlier models is the decision stage. We employ a maximum-likelihood approach based on the Fisher information formalism instead of the Minkowski norm approximation used in earlier studies.^{9,57,58,68} This approximation assumes that threshold is reached when the Minkowski norm of the differences in the mean responses to two alternative stimuli (expressed as multiples of the square root of the response variance) reaches unity (Subsection 2.B.5). Typically, the Minkowski norm is computed with an exponent Q between two⁵⁸ and four.^{11,24,57,58,68} To assess the importance of the decision stage for the quality of the model predictions, we fit the data of three observers using the Minkowski norm with $Q = 3$ as a decision stage. Figure 7 and Table 1 show representative results for observer SC.

Although the overall quality of the fit is only slightly worse for the Minkowski norm, the Minkowski norm consistently fails to predict the qualitative pattern of results in Exps. 4 and 5. Specifically, it predicts a nonmonotonic dependence of threshold elevation on mask contrast (Exp. 4) and too little threshold elevation for masks of different spatial frequencies (Exp. 5). In addition, the Minkowski norm consistently predicts a narrower filter tuning in spatial frequency (σ_λ of 0.13, 0.35, and 0.43 oct. for three observers) than is biologically plausible.

The inferior predictive power of the Minkowski decision reflects a fundamental problem that relates to its differential treatment of contrast discrimination and masking thresholds, on the one hand, and of orientation and spatial frequency discrimination thresholds, on the other hand. As illustrated in Fig. 3, contrast information is

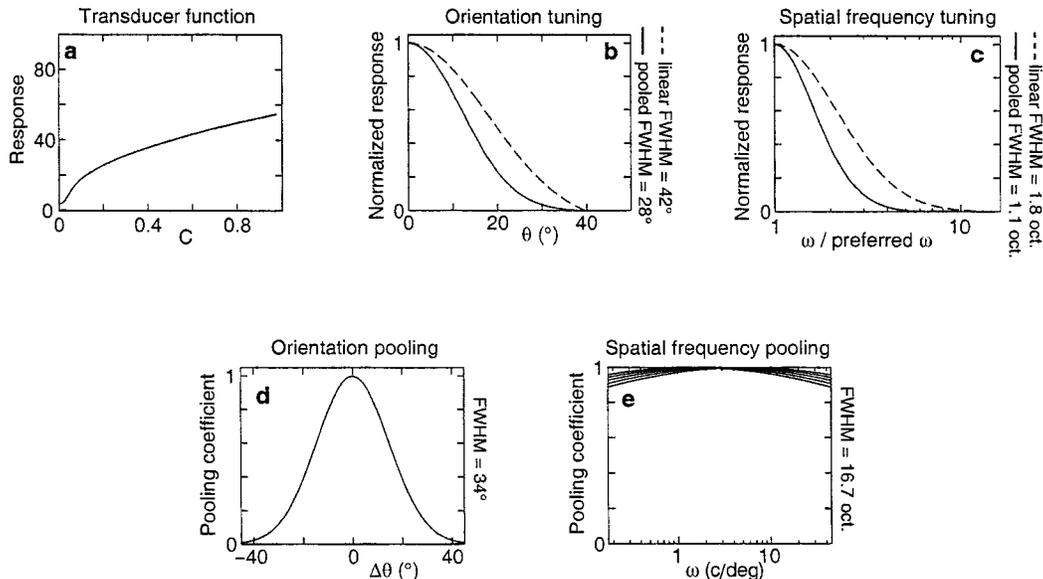


Fig. 6. Functional properties of the optimal model for observer SC. (a) The effective contrast response function exhibits the sigmoidal shape postulated by most psychophysical models. (b) and (c) The effective tuning functions for orientation and spatial frequency are approximately Gaussian but are 30%–40% narrower than the original tuning functions (dashed curves). (d) and (e) Relative weights with which different filters contribute to divisive inhibition. Inhibition derives from filters tuned to similar orientations (difference less than 40°). The range of spatial frequencies contributing to divisive inhibition is broad but only poorly constrained by the data.

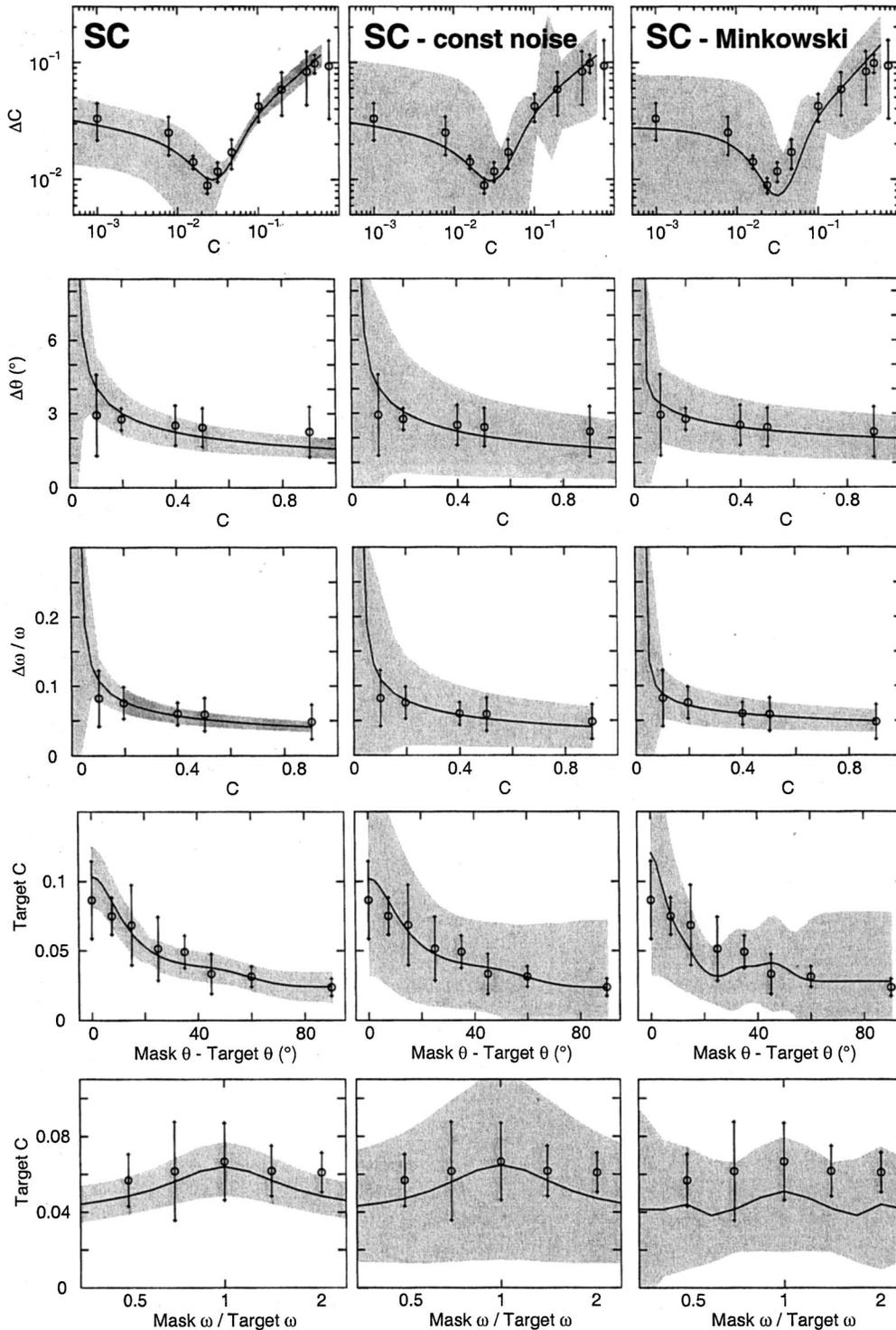


Fig. 7. Measured and predicted thresholds for two variants of the model (observer SC). The first column shows the fit of the standard model, for comparison. The second column shows the fit of a model variant using flat noise instead of proportional noise. The quality of the fit is comparable to that of the standard model, except in that it is less robust with respect to contrast-masking experiments (as indicated by relatively broad gray regions). The third column shows a model variant using a suboptimal decision based on the Minkowski norm. The fit is inferior to that of the standard model, in particular with respect to the contrast-masking experiments. Thus the statistically efficient decision contributes significantly to the success of the standard model.

concentrated in one, but orientation and spatial frequency information in two, filter subpopulations. The Minkowski decision differs from the Fisher information decision in how information from the two subpopulations

is combined: Whereas the Minkowski discriminability increases by a factor of $2^{1/Q}$, the Fisher discriminability increases by a factor of $2^{1/2}$ (Appendix B). As a result, the Minkowski decision with $Q > 2$ either overestimates

orientation and spatial frequency thresholds or underestimates contrast and masking thresholds, depending on how the overall sensitivity of the model is set. Thus the Minkowski decision contributes to the problems encountered by previous models in accounting simultaneously for different types of thresholds.⁶⁸

D. Contrast Dependence of Thresholds: a Reply to Bowne

In an influential paper, Bowne pointed out that a large class of models cannot explain the differential contrast dependence that the thresholds of human observers typically exhibit in the high-contrast regime.⁶⁸ For most observers, relative thresholds for contrast improve substantially with stimulus contrast ($\Delta C/C \propto C^{-0.3}$), whereas thresholds for orientation, spatial frequency, and other attributes improve little or not at all (e.g., $\Delta \theta \propto C^{-0.1}$). Contrary to this observation, many filter-based models predict that the contrast dependence of all thresholds should be the same (see also Appendix B).

As our model accurately predicts the contrast dependence of all investigated thresholds, we wished to understand where Bowne's seemingly general argument fails. To this end, we manually adjusted model parameters to obtain (i) Guilford's law⁹⁹ for contrast thresholds ($\Delta C \propto C^{0.75}$) and (ii) no contrast dependence for orientation thresholds ($\Delta \theta \propto \text{constant}$) [Figs. 8(a) and 8(b)]. The key for obtaining this differential contrast dependence turns out to be the exact shape of the sigmoidal contrast response function. In our model, this shape distorts the orientation tuning curves in a contrast-dependent manner [Figs. 8(c) and 8(d)]. In the region that determines orientation thresholds (i.e., $\pm 15^\circ$), the slope of the tuning function increases less with contrast than the height [as illustrated in Fig. 8(c)], and this shortfall suffices to keep orientation thresholds constant.

To see this point, consider the dependence of the Fisher information on the slope and height of the tuning function:

$$J_{\lambda, \theta} \approx \frac{(\partial R_{\lambda, \theta} / \partial \theta)^2}{R_{\lambda, \theta}}. \quad (10)$$

For the most informative units ($\theta = \pm 15^\circ$), the height $R_{\lambda, \theta}$ increases approximately threefold between contrasts 0.2 and 0.9, whereas the slope ($\partial R_{\lambda, \theta} / \partial \theta$) increases only approximately 1.75-fold. As a result of the lower increase of the slope, $R_{\lambda, \theta}$ and $(\partial R_{\lambda, \theta} / \partial \theta)^2$ increase by the same factor, and the Fisher information remains the same. This explains why orientation (and, similarly, spatial frequency) thresholds exhibit so little contrast dependence in the high-contrast regime. Note that an analogous argument can be made for the signal-to-noise ratio $\Delta R_{\lambda, \theta} / V_{\lambda, \theta}$, so that this conclusion holds irrespective of the decision stage used (i.e., Fisher information or Minkowski norm).

4. DISCUSSION

A. Common Basis of Spatial Vision

To ascertain whether different aspects of spatial vision reflect the same level of visual processing, we measured a

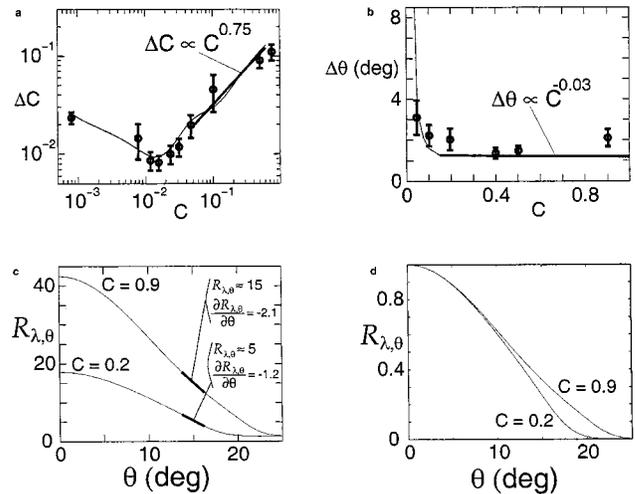


Fig. 8. Differential contrast dependence of thresholds. The model has been manually tuned such as to simultaneously predict contrast thresholds (a, Exp. 1) following Guilford's law ($\Delta C \propto C^{0.75}$), but contrast-independent orientation thresholds (b, Exp. 2) ($\Delta \theta \propto C^{-0.03}$). Looking at the internals of the model reveals that, although the unit responses, $R_{\lambda, \theta}$, increase for units of all orientations (c), the increase is more pronounced at the tails of the orientation tuning curve (d). As a result, the slope of the tuning curve at $\pm 15^\circ$ increases more slowly than its height. Specifically, as contrast increases from $C = 0.2$ to $C = 0.9$ curves, $R_{\lambda, \theta}$ increases by a factor of 3, whereas $\partial R_{\lambda, \theta} / \partial \theta$ increases only by a factor of 1.75 (for $\theta \pm 15^\circ$). As a result, the Fisher information with respect to contrast, which is approximately proportional to $(\partial R_{\lambda, \theta} / \partial \theta)^2 / R_{\lambda, \theta}$, does not increase with C .

wide range of thresholds with Gabor patterns presented at 4° of eccentricity. Previous studies have tended to focus either on contrast-increment and masking thresholds^{6,9,11,12,46} or on orientation and spatial frequency discrimination thresholds^{22,24,65,94,100,101} and thus were unable to address this larger question. The impetus for revisiting the basis of spatial vision at this time is provided by recent single-unit work in cat and monkey, which suggests that the behavioral thresholds in question may reflect neuronal response properties as early as primary visual cortex.^{20,63,102}

To test the hypothesis that spatial vision reflects a common neural basis, we employ a consensus model that combines components from several models of psychophysical^{9,42,42,45,58} and neuronal sensitivity^{63,86,103} to spatial patterns. Whenever possible, we substituted the more generic components of neural network models for the more specialized components customary in psychophysical models (e.g., Gaussian tuning functions, Poisson noise, ideal-observer decision). However, the basic architecture of our model (e.g., filter population, divisive inhibition) is fully consistent with previous psychophysical models. Our results demonstrate that a single model accounts quantitatively for all investigated thresholds and is fully consistent with a common basis of spatial vision.

We fit the ten parameters of our model to thirty-four threshold measurements from five separate experiments. The fitting procedure affords an effectively exhaustive search through the ten-dimensional parameter space and includes several randomly chosen starting points as well as two independent methods of iteration. The overall

quality of fit is good, and the residual error is generally comparable to the precision of measurement. To test the consistency of the result, we fit the model to separate data sets from three different observers. In spite of substantial differences in the threshold data, the best-fitting parameter values are almost always consistent and differ by 10% to 30% of their value between observers.

To assess robustness of fit, we compute a tolerance region around the optimal value of each parameter, within which the overall quality of the fit degrades by less than 5% (Appendix C). Almost all parameters are tightly constrained by data, some to within a few percent of their optimal value. The robustness of the fit is illustrated graphically by the gray regions in Figs. 5 and 7, which are generated by allowing all parameters to vary within their respective tolerance regions.

B. Relationship to Previous Models

As in many recent psychophysical models,^{19,45,57,58} we postulate visual filters tuned to a range of orientations and spatial frequencies but all centered at the same spatial location. To reduce the number of free parameters, we assume similar response characteristics for all filters (i.e., same contrast gain, background noise, width of orientation tuning, and width of spatial frequency tuning in octaves).

We chose visual filters defined by their polar response in the Fourier domain, which exhibit Gaussian tuning in orientation and spatial frequency. As functions of visual space, these filters closely resemble the multilobed functions employed by other models except that they are not separable along x and y dimensions. We saw no particular need to retain separability, as this property is unlikely to have any functional bearing. The advantage of filters with Gaussian tuning is that they render the effect of nonlinear interactions more transparent and easier to characterize.

We follow other recent models^{9,19,30,42–45,57,58} in normalizing visual filters responses by divisive inhibition.^{52,69} However, the parameters produced by our fitting procedure depart from other models in two notable respects: First, filter responses are taken to a relatively high exponent before being subjected to divisive inhibition (i.e., γ , $\delta \approx 3.0$ to 3.5). Previous models typically use values of γ , δ between 2.0 and 2.5 with a difference of 0.4 to 0.5.^{9,57,58} Second, only filters tuned to a relatively narrow range of orientations contribute to divisive inhibition. This near-orientation inhibition is consistent with the findings of a number of previous studies,^{12,104} but conflicts with the clear evidence for cross-orientation inhibition reported by other studies.^{9,105} As the latter group of studies used spatially extensive stimuli, this raises the possibility that cross-orientation inhibition may originate at more distant stimulus locations than near-orientation inhibition.

This difference in parameter values is of some functional importance. In our model, divisive inhibition leads to a substantial sharpening of the effective tuning for orientation and spatial frequency and also leads to a contrast response function of a particular sigmoidal shape. Together, these properties produce the differential contrast dependence of various types of thresholds

that is observed psychophysically (strong dependence of contrast-increment thresholds, weak dependence of orientation/spatial frequency discrimination thresholds). Other parameter choices would not have been able to account for this aspect of the psychophysical data.⁶⁸

Consistent with recent neural models,^{86,103} we assume that the variance of responses increases roughly in proportion to their means (proportional noise or Poisson^a noise). Psychophysical models typically assume noise of constant variance (flat noise).^{6,9,19,45,57,58} Although different noise assumptions are not interchangeable, in the present context proportional and flat noise afford essentially the same quality of fit. We use proportional noise to facilitate comparison with neural responses.

Our model incorporates an ideal-observer decision based on the Fisher-information framework.^{64,86,103} By relying on exact numerical calculations rather than analytical approximations, we have extended this framework to arbitrary psychophysical discriminations and sparse filter populations (Appendixes A and B). Previous psychophysical models have used a nonideal decision strategy based on the Minkowski norm of filter responses.^{45,57,58,68} An ideal decision strategy predicts our data significantly better than the nonideal strategy used by other models. For example, an ideal strategy predicts the relative levels of contrast-increment thresholds and orientation (spatial frequency) discrimination thresholds far better than does the nonideal strategy.

C. Relationship to Physiology

The response of neurons in primary visual cortex typically saturates at a given contrast and is best described by a hyperbolic ratio function.^{52,106,107} This is at variance with our model, where responses continue to increase with contrast (as a power function with an exponent of approximately 0.5). A likely reason for the discrepancy is that model responses reflect the average response of a diverse population of neurons that saturate at different contrasts.^{9,106}

The tuning width of neurons in primary visual cortex of macaque has been estimated to be $20 \pm 9^\circ$ for orientation and 0.76 ± 0.30 oct. for spatial frequency (half-width at half-maximum).^{17,20,65,76} This agrees reasonably well with the effective tuning half-widths of our model units, which are 14° for orientation and 0.56 oct. for spatial frequency, if one considers that psychophysical performance is likely to reflect the best-tuned neurons of a diverse population.^{94,100,108–110} We note also that tuning widths in primary visual cortex appear to be independent of contrast,⁷⁶ which is once again consistent with our model.

The variance of neuronal responses is roughly proportional to the mean, the exact relation being a power function with an exponent of 1.1 to 1.2.^{20,79,80} Our model agrees closely with these values, as the best-fitting exponent ranged between 1.0 and 1.3. Note, however, that our model assumes that responses vary independently of each other. This is not quite true in cortex, although the conditional covariance between visual cortical neurons (i.e., the correlation not due to the stimulus) is small.^{109,111} In our model, a population of units always contains more information than the most informative individual unit, because Fisher information is additive for

independent noise. In cortex, however, individual neurons may encode as much information about a stimulus as the animal as a whole.^{94,100,108–110} The reason for this discrepancy may lie in conditional covariance between neuronal responses, as the information of individual units would no longer be additive.

A particularly interesting issue is the neural basis of divisive inhibition. Although divisive inhibition was originally considered the result of shunting inhibition at the level of individual neurons,⁵³ it is now simply thought to provide a convenient description of the collective behavior of neural circuits in primary visual cortex.^{54,60,63,112,113} The function of these circuits remains controversial and may have to do with contrast adaptation, with the sharpening of neuronal tuning, or simply with gain control.^{52,60,114,115} The circuits that we describe in terms of divisive inhibition are likely to be found among short-range intrinsic connections in primary visual cortex.^{48,75,104,116,117} This is suggested by the functional organization of primary visual cortex and the fact that the circuits in question involve neurons with overlapping receptive fields and similar tuning properties. Recurrent excitation and inhibition within cortical columns are likely to play an important role as well.^{49,112,118} Indeed, detailed computational models combining short-range inhibition and recurrent excitation exhibit functionalities that are very similar to divisive inhibition (e.g., sharpening orientation tuning and reducing its dependence on contrast).^{59,60}

Another feature of the interactions implemented in the model was that we found the inhibitory pool to be narrow in the orientation domain. This property results from our masking data, which do not exhibit any cross-orientation inhibition. This connectivity is directly supported by studies of intrinsic connections in cat hypercolumns (e.g., Ref. 49), where excitatory connections are found predominantly between cells with the same tuning, and inhibitory connections are found between cells with similar tuning but not between cells with orthogonal tuning (see also Refs. 119 and 120). Our model does not make a clear prediction for either narrow^{52,121} or broad^{56,122} pools in spatial period.

Although our decision stage is purely abstract and is not intended as a model of any particular level of cortical processing, it is interesting to note that neural networks are well suited for computing statistically efficient estimates of stimulus attributes such as contrast, orientation, and spatial frequency.^{64,85,86,103}

5. CONCLUSIONS

We conclude that a wide range of spatial vision thresholds reflect a single level of visual processing, most likely corresponding to primary visual cortex. Furthermore, the visual processing in question is described quite well by divisive inhibition among overlapping visual filters. Both conclusions follow from our finding that divisive inhibition simultaneously predicts contrast-increment thresholds, orientation and spatial frequency discrimination thresholds, and contrast-masking thresholds. The parameters of divisive inhibition, as inferred from threshold

data, are tightly constrained. For the most part, they are in excellent agreement with what is known about visual processing at the level of primary visual cortex. To account for these behavioral thresholds, we found it important to employ a statistically efficient decision that avoids any bias in favor of one threshold or another. To this end, we describe a generalized Fisher-information approach that can be adapted to arbitrary psychophysical tasks (see Appendix A).

We have shown elsewhere that the present model accounts also for threshold changes by visual attention.^{70,123,124} Specifically, a change in one parameter of divisive inhibition (the exponent of the power law) replicates the contrast-increment thresholds, orientation, and spatial frequency discrimination thresholds, as well as contrast-masking thresholds for stimuli that are poorly attended. On the basis of these findings, we have proposed that divisive inhibition implements a winner-take-all competition among visual filters, a competition that is intensified by visual attention. By intensifying competition, attention enhances visual filters that respond relatively well to a given stimulus and attenuates visual filters that respond relatively less well. In the (near) absence of competition, the system acts in a more linear, superpositional mode. Thus divisive inhibition provides a surprisingly general framework for analyzing the perceptual and neural basis for a wide range of basic visual thresholds. The main challenges lying ahead are to link divisive inhibition to specific circuits in visual cortex^{59,112} and to characterize the important inhibitory and excitatory influences originating at more distant visual locations.^{32,33,35,36}

APPENDIX A: FISHER INFORMATION FOR A GAUSSIAN RANDOM VARIABLE

Consider a Gaussian random variable X with mean μ and variance σ^2 , both of which are functions of a stimulus attribute ζ . We now derive the Fisher information of X with respect to ζ . The probability of observing x given ζ , $p(x|\zeta)$, and its derivative with respect to ζ , $p'(x|\zeta)$, are

$$p(x|\zeta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad (\text{A1})$$

$$p'(x|\zeta) = p(x|\zeta) \left[\frac{(x-\mu)^2}{\sigma^3} \sigma' + \frac{x-\mu}{\sigma^2} \mu' - \frac{\sigma'}{\sigma} \right]. \quad (\text{A2})$$

If we denote the expectation of $f(x)$ with respect to $p(x|\zeta)$ with $E[f(x)]$, then, for any function f ,

$$E[f(x)] = \int_{-\infty}^{+\infty} f(x)p(x|\zeta)dx, \quad (\text{A3})$$

the definition of Fisher information (e.g., Ref. 83) yields

$$\begin{aligned}
J(\zeta) &= E \left[\left(\frac{\partial}{\partial \zeta} \log p(X|\zeta) \right)^2 \right] \\
&= E \left[\frac{p'(X|\zeta)^2}{p(X|\zeta)^2} \right] \\
&= E \left[\left(\frac{(X - \mu)^2}{\sigma^3} \sigma' + \frac{X - \mu}{\sigma^2} \mu' - \frac{\sigma'}{\sigma} \right)^2 \right] \\
&= \frac{1}{\sigma^6} [(\mu^2 \sigma^2 \mu'^2 - 2\mu^3 \sigma \mu' \sigma' + 2\mu \sigma^3 \mu' \sigma' \\
&\quad + \mu^4 \sigma'^2 - 2\mu^2 \sigma^2 \sigma'^2 + \sigma^4 \sigma'^2) \\
&\quad + (-2\mu \sigma^2 \mu'^2 + 6\mu^2 \sigma \mu' \sigma' - 2\sigma^3 \mu' \sigma' \\
&\quad - 4\mu^3 \sigma'^2 + 4\mu \sigma^2 \sigma'^2) E[X] + (\sigma^2 \mu'^2 \\
&\quad - 6\mu \sigma \mu' \sigma' + 6\mu^2 \sigma'^2 - 2\sigma^2 \sigma'^2) E[X^2] \\
&\quad + (2\sigma \mu' \sigma' - 4\mu \sigma'^2) E[X^3] + (\sigma'^2) E[X^4]] \\
&= \frac{\mu'^2 + 2\sigma'^2}{\sigma^2}, \tag{A4}
\end{aligned}$$

where μ' and σ' are the derivatives of μ and σ , respectively, with respect to ζ . In the special case of $\sigma^2 = \mu^\alpha$ (Poisson $^\alpha$ noise), the Fisher information becomes

$$J(\zeta) = \frac{\mu'^2}{\mu^2} \left(\mu^{2-\alpha} + \frac{\alpha^2}{2} \right). \tag{A5}$$

This expression is identical to Eq. (7).

APPENDIX B: COMPARISON OF NOISE AND DECISION MODELS

It is instructive to compare different noise and decision models for filters with relatively simple response properties for which the different approaches yield similar, and in some cases identical, results. Of course, this close correspondence breaks down for filters with more complex response properties, such as the sigmoidal contrast-response functions used in our model. First we derive the Fisher information for Poisson and constant noise, and then we compare decisions based on one filter, on all filters, and on the Minkowski norm.

1. Poisson Noise

Consider a visual filter i tuned to orientation θ_i and spatial frequency ω_i whose response is a Gaussian random variable with mean R_i and variance V_i^2 given by

$$\begin{aligned}
R_i &= A c^f \exp \left[-\frac{(\theta - \theta_i)^2}{2\sigma_\theta^2} \right] \exp \left[-\frac{(\omega - \omega_i)^2}{2\sigma_\omega^2} \right], \\
V_i^2 &= \beta R_i, \tag{B1}
\end{aligned}$$

where A is the sensitivity; f is the power of the contrast dependence; σ_θ and σ_ω are the tuning widths for orientation and spatial frequency, respectively; β the noise level; and c , θ , and ω are the contrast, orientation, and spatial frequency of the stimulus, respectively. Note that such a filter differs from those used in our model by neglecting

background activity, by using a power function for the contrast dependence, by fixing $\alpha = 1$, and by being independent of other filters (i.e., filters with other tuning properties).

To derive the Fisher information with respect to c , θ , and ω , we note that

$$\begin{aligned}
\frac{\partial R_i}{\partial c} &= \frac{f R_i}{c}, & \frac{\partial R_i}{\partial \theta} &= -R_i \frac{\theta - \theta_i}{\sigma_\theta^2}, \\
\frac{\partial R_i}{\partial \omega} &= -R_i \frac{\omega - \omega_i}{\sigma_\omega^2}. \tag{B2}
\end{aligned}$$

Equation (7) yields the Fisher information with respect to attribute ζ , J_i^ζ , where $\zeta = c, \theta, \omega$:

$$\begin{aligned}
J_i^\zeta &= \frac{1}{R_i^2} \left(\frac{\partial R_i}{\partial \zeta} \right)^2 \left(\frac{R_i}{\beta} + \frac{1}{2} \right) \approx \frac{1}{R_i^2} \left(\frac{\partial R_i}{\partial \zeta} \right)^2 \frac{R_i}{\beta}, \\
J_i^c &\approx \frac{f^2}{c^2} \frac{R_i}{\beta}, & J_i^\theta &\approx \frac{(\theta - \theta_i)^2}{\sigma_\theta^4} \frac{R_i}{\beta}, \\
J_i^\omega &\approx \frac{(\omega - \omega_i)^2}{\sigma_\omega^4} \frac{R_i}{\beta}. \tag{B3}
\end{aligned}$$

2. Constant Noise

Identical results can be obtained for constant noise, at least for the simple visual filters considered here. Let

$$\begin{aligned}
R_i &= \sqrt{A} c^{f/2} \exp \left[-\frac{(\theta - \theta_i)^2}{4\sigma_\theta^2} \right] \exp \left[-\frac{(\omega - \omega_i)^2}{4\sigma_\omega^2} \right], \\
V_i^2 &= \frac{\beta}{4}, \tag{B4}
\end{aligned}$$

where the constants have been chosen to facilitate comparison with the case of Poisson noise. Note that the signal-to-noise ratio R_i/V_i is four times larger with constant than with Poisson noise. The derivatives for constant noise are

$$\begin{aligned}
\frac{\partial R_i}{\partial c} &= \frac{f R_i}{2c}, & \frac{\partial R_i}{\partial \theta} &= -R_i \frac{\theta - \theta_i}{2\sigma_\theta^2}, \\
\frac{\partial R_i}{\partial \omega} &= -R_i \frac{\omega - \omega_i}{2\sigma_\omega^2}, \tag{B5}
\end{aligned}$$

and the Fisher information with respect to attribute ζ is

$$\begin{aligned}
J_i^\zeta &= \frac{4}{\beta} \left(\frac{\partial R_i}{\partial \zeta} \right)^2, & J_i^c &= \frac{f^2}{c^2} \frac{R_i^2}{\beta}, \\
J_i^\theta &= \frac{(\theta - \theta_i)^2}{\sigma_\theta^4} \frac{R_i^2}{\beta}, & J_i^\omega &= \frac{(\omega - \omega_i)^2}{\sigma_\omega^4} \frac{R_i^2}{\beta}. \tag{B6}
\end{aligned}$$

This result is identical to the one obtained for Poisson noise.

3. Most Informative Filters

The threshold in a 2AFC experiment can be related to the Fisher information by letting performance equal 3/4 and reformulating Eq. (6) to

$$\Delta\zeta = |\zeta_1 - \zeta_2| = k \sqrt{\frac{1/J(\zeta_1) + 1/J(\zeta_2)}{2}}$$

$$\approx \frac{k}{\sqrt{J[(\zeta_1 + \zeta_2)/2]}}; \quad k = 2 \operatorname{erf}^{-1} \frac{1}{2}. \quad (\text{B7})$$

If we base the decision on only one filter, we can choose the most informative filters, in other words, the ones with the largest Fisher information. For contrast discrimination, J_i^c is maximal when $\theta_i = \theta$ and $\omega_i = \omega$; for orientation discrimination, J_i^θ is maximal when $(\theta_i - \theta)^2 = 2\sigma_\theta^2$ and $\omega_i = \omega$; and for spatial frequency discrimination, J_i^ω is maximal when $(\omega_i - \omega)^2 = 2\sigma_\omega^2$ and $\theta_i = \theta$. The thresholds that result from a decision based on these filters can be obtained from Eq. (B3) or Eq. (B6):

$$\frac{\Delta c}{c} \approx 2k \sqrt{\frac{\beta}{Af^2 c^f}}, \quad \Delta\theta \approx \sqrt{2ek} \sqrt{\frac{\beta\sigma_\theta^2}{Ac^f}},$$

$$\Delta\omega \approx \sqrt{2ek} \sqrt{\frac{\beta\sigma_\omega^2}{Ac^f}}. \quad (\text{B8})$$

4. All Filters

Alternatively, we may base the decision on all filters by computing the Fisher information for the entire population. For simplicity, consider a population of filters spaced $\Delta\theta$ and $\Delta\omega$ apart and covering the entire orientation/spatial frequency plane. The total Fisher information for this population is

$$J_{\text{tot}} = \sum_{\theta_i, \omega_i} J_i = \frac{1}{\Delta\theta\Delta\omega} \sum_{\theta_i, \omega_i} J_i \Delta\theta\Delta\omega \approx \frac{\rho}{\sigma_\theta\sigma_\omega} \iint J_i d\theta d\omega, \quad (\text{B9})$$

where ρ is the filter density in units of $1/\sigma_\theta\sigma_\omega$. As the filter density increases, the approximation becomes more and more exact. With the help of

$$\frac{1}{\sqrt{2\pi\sigma_\theta^2}} \int \exp\left[-\frac{(\theta - \theta_i)^2}{2\sigma_\theta^2}\right] d\theta = 1,$$

$$\frac{1}{\sqrt{2\pi\sigma_\theta^2}} \int \frac{(\theta - \theta_i)^2}{2\sigma_\theta^2} \exp\left[-\frac{(\theta - \theta_i)^2}{2\sigma_\theta^2}\right] d\theta = \frac{1}{2},$$

we obtain the following expressions for the total Fisher information and the discrimination threshold:

$$J_{\text{tot}}^c \approx \frac{\rho}{\sigma_\theta\sigma_\omega} \iint J_i^c d\theta d\omega = \frac{2\pi\sigma_\theta\sigma_\omega Af^2 c^f \rho}{\beta c^2},$$

$$\frac{\Delta c}{c} \approx \sqrt{\frac{2}{\pi}} k \sqrt{\frac{\beta}{Af^2 c^f \rho}}, \quad (\text{B10})$$

$$J_{\text{tot}}^\theta \approx \frac{\rho}{\sigma_\theta\sigma_\omega} \iint J_i^\theta d\theta d\omega = \frac{2\pi\sigma_\omega Ac^f \rho}{\beta\sigma_\theta},$$

$$\Delta\theta \approx \sqrt{\frac{2}{\pi}} k \sqrt{\frac{\beta\sigma_\theta^2}{Ac^f \rho}}, \quad (\text{B11})$$

$$J_{\text{tot}}^\omega \approx \frac{\rho}{\sigma_\theta\sigma_\omega} \iint J_i^\omega d\theta d\omega = \frac{2\pi\sigma_\theta Ac^f \rho}{\beta\sigma_\omega},$$

$$\Delta\omega \approx \sqrt{\frac{2}{\pi}} k \sqrt{\frac{\beta\sigma_\omega^2}{Ac^f \rho}}. \quad (\text{B12})$$

Note that the contrast threshold based on all filters is smaller than the threshold based on the most informative filter, by a factor of $1/\sqrt{2\pi}$. The difference in orientation and spatial frequency thresholds is even larger, by a factor of $1/\sqrt{e\pi}$.

5. Minkowski Norm

The Minkowski norm has been one of the most popular ways to model a perceptual decision.^{68,91} Because thresholds reflect the magnitude of the response difference with respect to two stimulus alternatives, one expects that thresholds will be proportional to the derivative of the response R_i with respect to a stimulus attribute ζ . Specifically, one may postulate that threshold is reached when the response difference ΔR_i equals the standard deviation V_i of the response:

$$\Delta R_i = \left| \frac{\partial R_i}{\partial \zeta} \right| \Delta\zeta = V_i, \quad \frac{\Delta R_i}{V_i} = \frac{1}{V_i} \left| \frac{\partial R_i}{\partial \zeta} \right| \Delta\zeta = 1. \quad (\text{B13})$$

When the decision is based on multiple filters, threshold is reached when the Minkowski norm of the ratios $\Delta R_i/V_i$ reaches unity:

$$1 = \left(\sum_i \left| \frac{\Delta R_i}{V_i} \right|^Q \right)^{1/Q}. \quad (\text{B14})$$

The ratios $\Delta R_i/V_i$ stand in a simple relation to the Fisher information, which is obtained both for constant noise and for Poisson noise:

$$\frac{\Delta R_i}{V_i} = \frac{1}{\sqrt{\beta}} \left| \frac{\partial R_i}{\partial \zeta} \right| \Delta\zeta = \sqrt{J_i^\zeta} \Delta\zeta. \quad (\text{B15})$$

In the special case of $Q = 2$, the Minkowski norm predicts the same thresholds as does Fisher information, except for a proportionality constant:

$$1 = \left[\sum_i \left(\frac{\Delta R_i}{V_i} \right)^2 \right]^{1/2} = \Delta\zeta \left(\sum_i J_i^\zeta \right)^{1/2}, \quad \Delta\zeta = \frac{1}{\sqrt{J_{\text{tot}}^\zeta}}. \quad (\text{B16})$$

In the general case ($Q > 2$), however, the predictions of the Minkowski norm differ from those of Fisher information. This is particularly true when information is distributed over many filters. To see this, consider a population of N filters with identical $\Delta R_i/V_i = \sqrt{J_i^\zeta} = \sqrt{J_\zeta}$:

$$\Delta\zeta_{\text{Minkowski}} = \left[\sum_i \left(\frac{\Delta R_i}{V_i} \right)^Q \right]^{-1/Q} = N^{-1/Q} (J_\zeta)^{-1/2}$$

$$\Delta\zeta_{\text{Fisher}} = \left(\sum_i J_i^\zeta \right)^{-1/2} = N^{-1/2} = (J_\zeta)^{-1/2}. \quad (\text{B17})$$

The factor by which the Minkowski norm overpredicts thresholds, $N^{1/2-1/Q}$, grows with increasing N if $Q > 2$.

APPENDIX C: ANALYSIS OF MODEL FITS

The robustness and stability of model fits can be analyzed with the help of two approximations: a linear approximation of the threshold function and a quadratic approximation of the fit error function. These approximations allow us to derive analytical expressions for (i) the region around the optimal value to which each parameter is constrained by data and (ii) the varying extents to which different data points constrain the model.

Let X be the vector of ten model parameters and X_0 be the best-fit value of these parameters. Each threshold prediction $t_i(X)$ for each data point i is linearized around X_0 as

$$t_i(X) = t_i(X_0) + J_{t_i}^T(X - X_0), \quad (C1)$$

where J_{t_i} is the Jacobian of t_i at X_0 . Similarly, the fit error to the data $e(X)$ is approximated by the second-order formulation:

$$e(X) = e(X_0) + J_e^T(X - X_0) + (X - X_0)^T H_e (X - X_0), \quad (C2)$$

where J_e is the Jacobian and H_e the Hessian of e at X_0 . At the minimum of e in X_0 , we know that $J_e = 0$ and H_e is symmetric and positive.

To determine the tolerance range within which model parameters are constrained by data, we need to project the subspace $e(X) \leq e(X_0) + \epsilon$ onto each parameter's axis (see Ref. 95, Chap. 15 and Fig. 9 below). This projection yields a measure of the maximum variation obtained for each parameter when all parameters are allowed to arbitrarily vary while the fit error is kept within ϵ of the best-fit error. We can write this projection problem as a constrained optimization problem: Denoting E_j the basis vector corresponding to the axis of parameter j , we want to extremize

$$E_j^T X \quad \text{subject to} \quad (X - X_0)^T H_e (X - X_0) \leq \epsilon. \quad (C3)$$

To carry out this extremization, we write the Lagrange multiplier for this problem:

$$E_j^T X - \lambda [(X - X_0)^T H_e (X - X_0) - \epsilon]. \quad (C4)$$

Differentiating this expression successively with respect to each parameter yields a system of equations that we can write in matrix form:

$$E_j - 2\lambda H_e (X - X_0) = 0. \quad (C5)$$

Using this equation and the expression for the constraint, we solve for λ and then for X :

$$X = X_0 \pm \sqrt{\frac{\epsilon}{E_j^T H_e^{-1} E_j}} H_e^{-1} E_j, \quad (C6)$$

hence

$$E_j^T X = E_j^T X_0 \pm \sqrt{\epsilon E_j^T H_e^{-1} E_j}, \quad (C7)$$

and the tolerance range for parameter j (given in Table 1) is defined as the difference between $E_j^T X_0$ and the maximum or minimum of $E_j^T X$:

$$\sigma_j = \sqrt{\epsilon E_j^T H_e^{-1} E_j}. \quad (C8)$$

To determine the relative importance of different data points in constraining the model, we need to investigate the extent of the variations in each threshold prediction $t_i(X)$ when X is allowed to arbitrarily vary while ensuring that $e(X) \leq e(X_0) + \epsilon$. Since H_e is symmetric and positive, we can transform, through diagonalization of H_e , this quadratic constraint into a simple spherical inequality: We first decompose H_e as $H_e = V D V^{-1}$, where D is diagonal and V is an orthogonal change-of-basis matrix. Because $H_e \geq 0$, all its eigenvalues are positive, such that

$$H_e = V D^{1/2} D^{1/2} V^{-1}. \quad (C9)$$

In addition, because V is orthogonal, $V^{-1} = V^T$. We can now rewrite the constraint as

$$(X - X_0)^T V D^{1/2} D^{1/2} V^T (X - X_0) \leq \epsilon, \quad (C10)$$

$$[D^{1/2} V^T (X - X_0)]^T [D^{1/2} V^T (X - X_0)] \leq \epsilon, \quad (C11)$$

$$W^T W \leq \epsilon, \quad \text{i.e.,} \quad \|W\|^2 \leq \epsilon, \quad W = D^{1/2} V^T (X - X_0). \quad (C12)$$

Consequently, the region of the parameter space where $e(X) < e(X_0) + \epsilon$ is the interior of the hypersphere $\|W\|^2 = \epsilon$. In order to compute the envelopes of all model predictions when the parameters are inside this hypersphere, we need to find the minimum and maximum of each $t_i(X)$ subject to the constraint $\|W\|^2 \leq \epsilon$. Since we locally approximate t_i by the linear form $t_i(X) = t_i(X_0) + J_{t_i}^T (X - X_0)$, we trivially know that the extrema of this linear form will be obtained for X such that $\|W\|^2 = \epsilon$; indeed, we simply write

$$t_i(X) = t_i(X_0) + J_{t_i}^T V^{-T} D^{-1/2} D^{1/2} V^T (X - X_0), \quad (C13)$$

$$t_i(X) = t_i(X_0) + K^T W \quad (C14)$$

with $K = D^{-1/2} V^T J_{t_i}$ such that the extremization problem becomes that of finding the minimum and maximum of

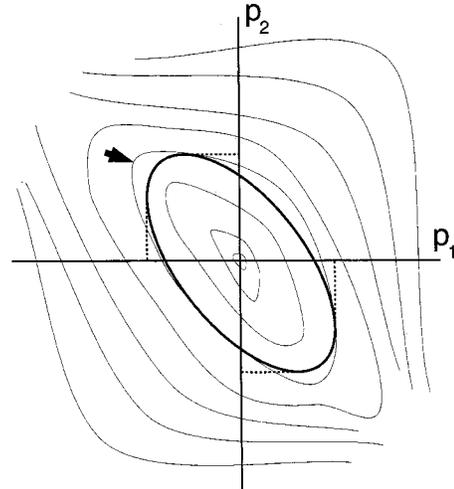


Fig. 9. Approximation of the error surface near the point of best fit, X_0 . The error surface is approximated by a paraboloid based on the Hessian matrix at X_0 . With this approximation, all points at which the fit error $e(X) \leq e(X_0) + \epsilon$ are within an ellipsoid (thick curve). The actual isocontour at which $e(X) \leq e(X_0) + \epsilon$ is indicated by the arrow. For each parameter, the tolerance range within which $e(X) \leq e(X_0) + \epsilon$ is obtained by projecting the ellipsoid onto the associated axis (dotted lines).

$t_i(X_0) + K^T W$ subject to $\|W\|^2 \leq \epsilon$. Because we have reduced the function to extremize to a simple dot product between two vectors, we now see that these extrema will be obtained when W is collinear with K and, more precisely, when

$$W = \pm \sqrt{\epsilon} \frac{K}{\|K\|}. \quad (\text{C15})$$

Finally, for the two values of W that extremize t_i , we can compute the minimum and the maximum of the model prediction for datapoint i from Eq. (C14):

$$t_i^{\text{extr}} = t_i(X_0) \pm \sqrt{\epsilon} \|K\|. \quad (\text{C16})$$

With this method, we can estimate the extrema of the range of threshold predictions obtained from arbitrary parameter values X within a certain neighborhood of the best-fit point X_0 [$e(X) \leq e(X_0) + \epsilon$]. The accuracy of this estimate is limited only by the accuracy of the (very reasonable) approximations for t_i and e .

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation-sponsored Neuromorphic Engineering Research Center, the National Institute of Mental Health, the U.S. Office of Naval Research, and the Sloan Center for Theoretical Neurobiology. We thank Amit Manwani for comments on various theoretical aspects and Pietro Perona for discussions and help with the experimental setup.

Corresponding author Christof Koch can be reached by e-mail at koch@klab.caltech.edu.

REFERENCES

- J. P. Thomas, "Model of the function of receptive fields in human vision," *Psychol. Rev.* **77**, 121–134 (1970).
- N. Graham and J. Nachmias, "Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channels models," *Vision Res.* **11**, 251–259 (1971).
- M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial-frequency channels in human vision," *J. Opt. Soc. Am.* **61**, 1176–1186 (1971).
- J. Nachmias and R. V. Sansbury, "Letter: grating contrast: discrimination may be better than detection," *Vision Res.* **14**, 1039–1042 (1974).
- C. F. Stromeyer and S. Klein, "Spatial frequency channels in human vision as asymmetric (edge) mechanisms," *Vision Res.* **14**, 1409–1420 (1974).
- G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Am.* **70**, 1458–1471 (1980).
- H. R. Wilson and J. R. Bergen, "A four mechanism model for threshold spatial vision," *Vision Res.* **19**, 19–32 (1979).
- H. R. Wilson, "Nonlinear processes in visual pattern discrimination," *Proc. Natl. Acad. Sci. USA* **90**, 9785–9790 (1993).
- J. M. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," *J. Opt. Soc. Am. A* **11**, 1710–1719 (1994).
- C. Blakemore and F. W. Campbell, "On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images," *J. Physiol. (London)* **203**, 237–260 (1969).
- H. R. Wilson, D. K. McFarlane, and G. C. Phillips, "Spatial frequency tuning of orientation selective units estimated by oblique masking," *Vision Res.* **23**, 873–882 (1983).
- G. C. Phillips and H. R. Wilson, "Orientation bandwidths of spatial mechanisms measured by masking," *J. Opt. Soc. Am. A* **1**, 226–232 (1984).
- C. F. Stromeyer and S. Klein, "Evidence against narrow-band spatial frequency channels in human vision: the detectability of frequency modulated gratings," *Vision Res.* **15**, 899–910 (1975).
- N. Graham and J. G. Robson, "Summation of very close spatial frequencies: the importance of spatial probability summation," *Vision Res.* **27**, 1997–2007 (1987).
- D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol. (London)* **195**, 215–243 (1968).
- L. Maffei and A. Fiorentini, "The visual cortex as a spatial frequency analyzer," *Vision Res.* **13**, 1255–1267 (1973).
- R. L. DeValois, D. G. Albrecht, and L. G. Thorell, "Spatial-frequency selectivity of cells in macaque visual cortex," *Vision Res.* **22**, 545–559 (1982).
- R. L. DeValois and K. K. DeValois, *Spatial Vision* (Oxford U. Press, New York, 1988).
- H. R. Wilson, D. Levi, L. Maffei, J. Rovamo, and R. DeValois, "The perception of form: retina to striate cortex," in *Visual Perception: The Neurophysiological Foundations*, L. Spillmann and J. S. Werner, eds. (Academic, San Diego, Calif., 1990), pp. 231–272.
- W. S. Geisler and D. G. Albrecht, "Visual cortex neurons in monkeys and cats: detection, discrimination, and identification," *Visual Neurosci.* **14**, 897–919 (1997).
- J. Ross and H. D. Speed, "Contrast adaptation and contrast masking in human vision," *Proc. R. Soc. London Ser. B* **246**, 61–69 (1991).
- C. A. Burbeck and D. Regan, "Independence of orientation and size in spatial discriminations," *J. Opt. Soc. Am.* **73**, 1691–1694 (1983).
- A. Bradley and B. C. Skottun, "The effects of large orientation and spatial frequency differences on spatial discriminations," *Vision Res.* **24**, 1889–1896 (1984).
- H. R. Wilson and D. J. Gelb, "Modified line-element theory for spatial-frequency and width discrimination," *J. Opt. Soc. Am. A* **1**, 124–131 (1984).
- S. A. Klein and D. M. Levi, "Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation," *J. Opt. Soc. Am. A* **2**, 1170–1190 (1985).
- H. R. Wilson, "Responses of spatial mechanisms can explain hyperacuity," *Vision Res.* **26**, 453–469 (1986).
- H. R. Wilson, "Model of peripheral and amblyopic hyperacuity," *Vision Res.* **31**, 967–982 (1991).
- U. Polat and D. Sagi, "Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments," *Vision Res.* **33**, 993–999 (1993).
- U. Polat and D. Sagi, "The architecture of perceptual spatial interactions," *Vision Res.* **34**, 73–78 (1994).
- F. Wilkinson, H. R. Wilson, and D. Ellemberg, "Lateral interactions in peripherally viewed texture arrays," *J. Opt. Soc. Am. A* **14**, 2057–2068 (1997).
- J. J. Knierim and D. C. van Essen, "Neuronal responses to static texture patterns in area V1 of the alert macaque monkey," *J. Neurophysiol.* **67**, 961–980 (1992).
- M. K. Kapadia, M. Ito, C. D. Gilbert, and G. Westheimer, "Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys," *Neuron* **15**, 843–856 (1995).
- A. M. Sillito, K. L. Grieve, H. E. Jones, J. Cudeiro, and J. Davis, "Visual cortical mechanisms detecting focal orientation discontinuities," *Nature* **378**, 492–496 (1995).
- U. Polat and A. M. Norcia, "Neurophysiological evidence for contrast dependent long-range facilitation and suppression in the human visual cortex," *Vision Res.* **36**, 2099–2109 (1996).

35. J. B. Levitt and J. S. Lund, "Contrast dependence of contextual effects in primate visual cortex," *Nature* **387**, 73–76 (1997).
36. U. Polat, K. Mizobe, M. W. Pettet, T. Kasamatsu, and A. M. Norcia, "Collinear stimuli regulate visual responses depending on cell's contrast threshold," *Nature* **391**, 580–584 (1998).
37. J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *J. Opt. Soc. Am. A* **7**, 923–932 (1990).
38. M. S. Landy and J. R. Bergen, "Texture segregation and orientation gradient," *Vision Res.* **31**, 679–691 (1991).
39. D. J. Field, A. Hayes, and R. F. Hess, "Contour integration by the human visual system: evidence for a local association field," *Vision Res.* **33**, 173–193 (1993).
40. M. B. Ben-Av and D. Sagi, "Perceptual grouping by similarity and proximity: experimental results can be predicted by intensity autocorrelations," *Vision Res.* **35**, 853–866 (1995).
41. M. Ito, G. Westheimer, and C. D. Gilbert, "Attention and perceptual learning modulate contextual influences on visual perception," *Neuron* **20**, 1191–1197 (1998).
42. H. R. Wilson and R. Humanski, "Spatial frequency adaptation and contrast gain control," *Vision Res.* **33**, 1133–1149 (1993).
43. B. Zenger and D. Sagi, "Isolating excitatory and inhibitory nonlinear spatial interactions involved in contrast detection," *Vision Res.* **36**, 2497–2513 (1996).
44. J. M. Foley and C. C. Chen, "Analysis of the effect of pattern adaptation on pattern pedestal effects: a two-process model," *Vision Res.* **37**, 2779–2788 (1997).
45. J. P. Thomas and L. A. Olzak, "Contrast gain control and fine spatial discriminations," *J. Opt. Soc. Am. A* **14**, 2392–2405 (1997).
46. H. R. Wilson, "A transducer function for threshold and suprathreshold human vision," *Biol. Cybern.* **38**, 171–178 (1980).
47. T. B. Lawton and C. W. Tyler, "On the role of X and simple cells in human contrast processing," *Vision Res.* **34**, 659–667 (1994).
48. K. Toyama, M. Kimura, and K. Tanaka, "Cross-correlation analysis of interneuronal connectivity in cat visual cortex," *J. Neurophysiol.* **46**, 191–201 (1981).
49. Y. Hata, T. Tsumoto, H. Sato, and H. Tamura, "Horizontal interactions between visual cortical neurones studied by cross-correlation analysis in the cat," *J. Physiol. (London)* **441**, 593–614 (1991).
50. A. B. Bonds, "Temporal dynamics of contrast gain in single cells of the cat striate cortex," *Visual Neurosci.* **6**, 239–255 (1991).
51. W. S. Geisler and D. G. Albrecht, "Cortical neurons: isolation of contrast gain control," *Vision Res.* **32**, 1409–1410 (1992).
52. D. J. Heeger, "Normalization of cell responses in cat striate cortex," *Visual Neurosci.* **9**, 181–197 (1992).
53. M. Carandini and D. J. Heeger, "Summation and division by neurons in primate visual cortex," *Science* **264**, 1333–1336 (1994).
54. B. Ahmed, J. D. Allison, R. J. Douglas, and K. A. Martin, "An intracellular study of the contrast-dependence of neuronal activity in cat visual cortex," *Cereb. Cortex* **7**, 559–570 (1997).
55. I. A. Shevelev, R. V. Novikova, N. A. Lazareva, A. S. Tikhomirov, and G. A. Sharaev, "Sensitivity to cross-like figures in the cat striate neurons," *Neuroscience* **69**, 51–57 (1995).
56. G. C. DeAngelis, J. G. Robson, I. Ohzawa, and R. D. Freeman, "Organization of suppression in receptive fields of neurons in cat visual cortex," *J. Neurophysiol.* **68**, 144–163 (1992).
57. P. C. Teo and D. J. Heeger, "Perceptual image distortion," in *Human Vision, Visual Processing, and Digital Display V*, B. E. Rogowitz and J. P. Allebach, eds., *Proc. SPIE* **2179**, 127–129 (1995).
58. A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Am. A* **14**, 2379–2391 (1997).
59. M. Carandini and D. L. Ringach, "Predictions of a recurrent model of orientation selectivity," *Vision Res.* **37**, 3061–3071 (1997).
60. D. C. Somers, S. B. Nelson, and M. Sur, "An emergent model of orientation selectivity in cat visual cortical simple cells," *J. Neurosci.* **15**, 5448–5465 (1995).
61. P. Adorjan, J. B. Levitt, J. S. Lund, and K. Obermayer, "A model for the intracortical origin of orientation preference and tuning in macaque striate cortex," *Visual Neurosci.* **16**, 303–318 (1999).
62. D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol. (London)* **160**, 106–154 (1962).
63. M. Carandini, D. J. Heeger, and J. A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *J. Neurosci.* **17**, 8621–8644 (1997).
64. A. Pouget, K. Zhang, S. Deneve, and P. E. Latham, "Statistically efficient estimation using population coding," *Neural Comput.* **10**, 373–401 (1998).
65. J. Hirsch and R. Hylton, "Limits of spatial-frequency discrimination as evidence of neural interpolation," *J. Opt. Soc. Am.* **72**, 1367–1374 (1982).
66. J. P. Thomas, "Underlying psychometric function for detecting gratings and identifying spatial frequency," *J. Opt. Soc. Am.* **73**, 751–758 (1983).
67. E. S. Richter and D. Yager, "Spatial-frequency difference thresholds for central and peripheral viewing," *J. Opt. Soc. Am. A* **1**, 1136–1139 (1984).
68. S. F. Bowne, "Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination," *Vision Res.* **30**, 449–461 (1990).
69. D. G. Albrecht and W. S. Geisler, "Motion selectivity and the contrast-response function of simple cells in the visual cortex," *Visual Neurosci.* **7**, 531–546 (1991).
70. D. K. Lee, L. Itti, C. Koch, and J. Braun, "Attention activates winner-take-all competition among visual filters," *Nat. Neurosci.* **2**, 375–381 (1999).
71. C. W. Tyler, "Color bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis," *Spatial Vis.* **10**, 369–377 (1997).
72. A. B. Watson, "Neural contrast sensitivity," in *Computational Models of Visual Processing*, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1990), pp. 95–108.
73. W. A. Weibull, "A statistical distribution function of wide applicability," *J. Appl. Mech.* **18**, 292–297 (1951).
74. A. Cowey and E. T. Rolls, "Human cortical magnification factor and its relation to visual acuity," *Exp. Brain Res.* **21**, 447–454 (1974).
75. G. G. Blasdel and G. Salama, "Voltage-sensitive dyes reveal a modular organization in monkey striate cortex," *Nature* **321**, 579–585 (1986).
76. B. C. Skottun, A. Bradley, G. Sclar, I. Ohzawa, and R. D. Freeman, "The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior," *J. Neurophysiol.* **57**, 773–786 (1987).
77. D. A. Pollen and S. F. Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science* **212**, 1409–1411 (1981).
78. M. Carandini, J. A. Movshon, and D. Ferster, "Pattern adaptation and cross-orientation interactions in the primary visual cortex," *Neuropharmacology* **37**, 501–511 (1998).
79. W. R. Softky and C. Koch, "The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs," *J. Neurosci.* **13**, 334–350 (1993).
80. M. C. Teich, R. G. Turcott, and R. M. Siegel, "Temporal correlation in cat striate-cortex neural spike trains," *IEEE Eng. Med. Biol. Mag.* Sept.–Oct. 1996, pp. 79–87.
81. D. M. Green and J. A. Swets, *Signal Detectability and Psychophysics* (Wiley, New York, 1966).
82. L. L. Scharf, *Statistical Signal Processing: Detection, Estimation and Time-Series Analysis* (Addison-Wesley, Reading, Mass., 1991).

83. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
84. H. P. Snippe and J. J. Koenderink, "Information in channel-coded systems: correlated receivers," *Biol. Cybern.* **67**, 183–190 (1992).
85. G. Mato and H. Sompolinsky, "Neural network models of perceptual learning of angle discrimination," *Neural Comput.* **8**, 270–299 (1996).
86. H. S. Seung and H. Sompolinsky, "Simple models for reading neuronal population codes," *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
87. S. Magnussen, M. W. Greenlee, and J. P. Thomas, "Parallel processing in visual short-term memory," *J. Exp. Psychol. Human Percept. Perform.* **22**, 202–212 (1996).
88. H. Dai, Q. Nguyen, and D. M. Green, "Decision rules of listeners in spectral-shape discrimination with or without signal-frequency uncertainty," *J. Acoust. Soc. Am.* **99**, 2298–2306 (1996).
89. H. Dai, "Signal-frequency uncertainty in spectral-shape discrimination: psychometric functions," *J. Acoust. Soc. Am.* **96**, 1388–1396 (1994).
90. P. Verghese and L. S. Stone, "Combining speed information across space," *Vision Res.* **35**, 2811–2823 (1995).
91. R. F. Quick, "A vector-magnitude model of contrast detection," *Kybernetik* **16**, 1299–1302 (1974).
92. M. W. Greenlee, "Spatial frequency discrimination of band-limited periodic targets: effects of stimulus contrast, bandwidth and retinal eccentricity," *Vision Res.* **32**, 275–283 (1992).
93. L. N. Thibos, D. L. Still, and A. Bradley, "Characterization of spatial aliasing and contrast sensitivity in peripheral vision," *Vision Res.* **36**, 249–258 (1996).
94. R. Vogels and G. A. Orban, "How well do response changes of striate neurons signal differences in orientation: a study in the discriminating monkey," *J. Neurosci.* **10**, 3543–3558 (1990).
95. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. (Cambridge U. Press, Cambridge, UK, 1992).
96. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.* **21**, 1087–1092 (1953).
97. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science* **220**, 671–680 (1983).
98. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 721–741 (1984).
99. J. P. Guilford, *Psychometric Methods* (McGraw-Hill, New York, 1954).
100. A. Bradley, B. C. Skottun, I. Ohzawa, G. Sclar, and R. D. Freeman, "Neurophysiological evaluation of the differential response model for orientation and spatial-frequency discrimination," *J. Opt. Soc. Am. A* **2**, 1607–1610 (1985).
101. M. A. Garcia-Perez and V. Sierra-Vazquez, "Do channels shift their tuning towards lower spatial frequencies in the periphery?" *Vision Res.* **36**, 3339–3372 (1996).
102. H. Sompolinsky and R. Shapley, "New perspectives on the mechanisms for orientation selectivity," *Curr. Opin. Neurobiol.* **7**, 514–522 (1997).
103. S. Deneve, P. E. Latham, and A. Pouget, "Reading population codes: a neural implementation of ideal observers [in process citation]," *Nat. Neurosci.* **2**, 740–745 (1999).
104. J. M. Crook, Z. F. Kisvarday, and U. T. Eysel, "GABA-induced inactivation of functionally characterized sites in cat striate cortex: effects on orientation tuning and direction selectivity," *Visual Neurosci.* **14**, 141–158 (1997).
105. L. A. Olzak and J. P. Thomas, "When orthogonal orientations are not processed independently," *Vision Res.* **31**, 51–57 (1991).
106. D. G. Albrecht and D. B. Hamilton, "Striate cortex of monkey and cat: contrast response function," *J. Neurophysiol.* **48**, 217–237 (1982).
107. W. S. Geisler and D. G. Albrecht, "Bayesian analysis of identification performance in monkey visual cortex: nonlinear mechanisms and stimulus certainty," *Vision Res.* **35**, 2723–2730 (1995).
108. K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon, "The analysis of visual motion: a comparison of neuronal and psychophysical performance," *J. Neurosci.* **12**, 4745–4765 (1992).
109. E. Zohary, M. N. Shadlen, and W. T. Newsome, "Correlated neuronal discharge rate and its implications for psychophysical performance," *Nature* **370**, 140–143 (1994); erratum **371**, 358 (1994).
110. M. N. Shadlen, K. H. Britten, W. T. Newsome, and J. A. Movshon, "A computational analysis of the relationship between neuronal and behavioral responses to visual motion," *J. Neurosci.* **16**, 1486–1510 (1996).
111. T. J. Gawne, T. W. Kjaer, J. A. Hertz, and B. J. Richmond, "Adjacent visual cortical complex cells share about 20% of their stimulus-related information," *Cereb. Cortex* **6**, 482–489 (1996).
112. R. J. Douglas, C. Koch, M. Mahowald, K. A. Martin, and H. H. Suarez, "Recurrent excitation in neocortical circuits," *Science* **269**, 981–985 (1995).
113. G. R. Holt and C. Koch, "Shunting inhibition does not have a divisive effect on firing rates," *Neural Comput.* **9**, 1001–1013 (1997).
114. D. Ferster and C. Koch, "Neuronal connections underlying orientation selectivity in cat visual cortex," *Trends Neurosci.* **10**, 187–192 (1987).
115. J. Bolz, C. D. Gilbert, and T. N. Wiesel, "Pharmacological analysis of cortical circuitry," *Trends Neurosci.* **12**, 292–296 (1989).
116. L. C. Katz, C. D. Gilbert, and T. N. Wiesel, "Local circuits and ocular dominance columns in monkey striate cortex," *J. Neurosci.* **9**, 1389–1399 (1989).
117. A. Das and C. D. Gilbert, "Topography of contextual modulations mediated by short-range interactions in primary visual cortex," *Nature* **399**, 655–661 (1999).
118. R. J. Douglas and K. A. Martin, "A functional microcircuit for cat visual cortex," *J. Physiol. (London)* **440**, 735–769 (1991).
119. D. Ferster, "Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex," *J. Neurosci.* **6**, 1284–1301 (1986).
120. D. Ferster, "Spatially opponent excitation and inhibition in simple cells of the cat visual cortex," *J. Neurosci.* **8**, 1172–1180 (1988).
121. A. S. Ramoa, M. Shadlen, B. C. Skottun, and R. D. Freeman, "A comparison of inhibition in orientation and spatial frequency selectivity of cat visual cortex," *Nature* **321**, 237–239 (1986).
122. R. W. Bowen and H. R. Wilson, "A two-process analysis of pattern masking," *Vision Res.* **34**, 645–657 (1994).
123. D. K. Lee, C. Koch, and J. Braun, "Spatial vision thresholds in the near absence of attention," *Vision Res.* **37**, 2409–2418 (1997).
124. L. Itti, J. Braun, D. K. Lee, and C. Koch, "Attentional modulation of human pattern discrimination psychophysics reproduced by a quantitative model," in *Advances in Neural Information Processing Systems*, Vol. 11, M. S. Kearns, S. A. Solla, and D. A. Cohn, eds. (MIT Press, Cambridge, Mass., 1999, pp. 699–705).