A Neural Population Model for Visual Pattern Detection

Robbe L. T. Goris
New York University and KU Leuven

Tom Putzeys and Johan Wagemans
KU Leuven

Felix A. Wichmann
Eberhard Karls Universität Tübingen and Bernstein Center for Computational Neuroscience Tübingen and
Max Planck Institute for Intelligent Systems, Tübingen, Germany

Pattern detection is the bedrock of modern vision science. Nearly half a century ago, psychophysicists advocated a quantitative theoretical framework that connected visual pattern detection with its neurophysiological underpinnings. In this theory, neurons in primary visual cortex constitute linear and independent visual channels whose output is linked to choice behavior in detection tasks via simple read-out mechanisms. This model has proven remarkably successful in accounting for threshold vision. It is fundamentally at odds, however, with current knowledge about the neurophysiological underpinnings of pattern vision. In addition, the principles put forward in the model fail to generalize to suprathreshold vision or perceptual tasks other than detection. We propose an alternative theory of detection in which perceptual decisions develop from maximum-likelihood decoding of a neurophysiologically inspired model of population activity in primary visual cortex. We demonstrate that this theory explains a broad range of classic detection results. With a single set of parameters, our model can account for several summation, adaptation, and uncertainty effects, thereby offering a new theoretical interpretation for the vast psychophysical literature on pattern detection.

Keywords: visual detection, neural population model, summation experiments, pattern adaptation, extrinsic uncertainty

Visual pattern detection experiments measure the minimal contrast required to reliably discriminate a stimulus from a blank. In the late 1960s and early 1970s, psychophysicists studying pattern detection pioneered a theory of visual processing that revolutionized our understanding of the computations performed by the early stages of the visual system. Within the framework of linear systems analysis, John Robson, Fergus Campbell, Jacob Nachmias, Norma Graham, and many others sought to connect psychophysical measurements of visual behavior with the sensory coding performed by single cells in visual cortex. The success of this endeavor inspired countless psychophysicists, physiologists, and computer vision scientists and is arguably the bedrock of modern vision science. But as the field moved forward, several weak spots of the early theoretical framework have come to light. The connection with the neural underpinnings of pattern vision turned out to be looser than anticipated, and the principles put forward in the theory to link neural activity to perceptual judgments failed to generalize to perceptual tasks other than detection.

Despite the early successes, visual pattern detection is currently in an unenviable position. On the one hand, it is considered solved and therefore not studied actively anymore. On the other hand, detection is often assumed to be a unique behavioral paradigm with limited relevance for suprathreshold vision or sensory information processing in general. This article rejects both notions and makes an effort to integrate the vast psychophysical literature on pattern detection into contemporary neuroscience. To this end, we examine how well a selection of classic detection phenomena (see Table 1) can be explained in an observer model that is tightly connected to current insights in visual neuroscience.[1] Built on the early theory, visual detection is described in an information processing framework with an explicit encoding and decoding stage. The nonlinear encoding stage gives rise to a fragmentary and noisy stimulus representation that serves as input to a linear decoder that

Robbe L. T. Goris, Center for Neural Science, New York University, and Laboratory of Experimental Psychology, KU Leuven, Leuven, Belgium. Tom Putzeys and Johan Wagemans, Laboratory of Experimental Psychology, KU Leuven. Felix A. Wichmann, Faculty of Science, University of Tübingen; Bernstein Center for Computational Neuroscience Tübingen, Tübingen, Germany; and Max Planck Institute for Intelligent Systems, Tübingen, Germany.

Correspondence concerning this article should be addressed to Robbe L. T. Goris, Center for Neural Science, New York University, 4 Washington Place, Room 809, New York, NY 10003. E-mail: robbe.goris@nyu.edu

---

[1] Interested readers may contact the corresponding author to get the Matlab code used to implement and fit the model.

Table 1
Pattern Detection Results Modeled in This Article

Behavioral measurement	Original source	Included in fitting[a]
Psychometric function for detection	Legge (1978)	
Contrast sensitivity function	Campbell & Robson (1968)	✓
Summation of far-apart frequencies	Graham & Nachmias (1971)	✓
Summation of near and far frequencies	Sachs et al. (1971)	
Square wave sensitivity	Campbell & Robson (1968)	✓
Detection of frequency-modulated gratings	Stromeyer & Klein (1975)	
Pattern adaptation: adapting time	Blakemore & Campbell (1969)	✓
Pattern adaptation: adapting contrast	Tolhurst (1972)	✓
Pattern adaptation: adapting frequency	Blakemore & Campbell (1969)	✓
Pattern adaptation: spectral composition	Blakemore & Campbell (1969), Tolhurst (1972)	
Extrinsic uncertainty: spatial frequency	Davis (1981), Davis & Graham (1981), Davis et al. (1983)	
Contrast discrimination	Nachmias & Sansbury (1974)	

[a] Measurements were part of the data set used to optimize the model parameters.

aims to maximize task performance without having detailed knowledge of the full sensory population response distribution. We analyze predictions of this model for summation, adaptation, and uncertainty effects in visual pattern detection and show that it explains a broad range of behavioral results using a single set of parameters.

Behavioral Pattern Detection: Data, Theory, and Impact

Organisms and machines facing multiple visual tasks require an initial processing stage that provides a generic image description (Adelson & Bergen, 1991). Nearly half a century ago, psychophysicists studying pattern detection pioneered the hypothesis that the human early visual system performs measurements of basic image properties using parallel pathways tuned to particular visual features such as the size and orientation of spatial variations in luminance, thereby laying the foundation of our current understanding of pattern vision (see Figure 1).

In an elegant series of behavioral experiments, a group of Cambridge researchers demonstrated that contrast detection thresholds of gratings with sinusoidal luminance profiles are hardly affected by either adding a second low-contrast grating (Campbell & Robson, 1968; Graham & Nachmias, 1971) or adapting to a high-contrast grating (Blakemore & Campbell, 1969; Pantle & Sekuler, 1968), provided that both sinusoids are sufficiently different in spatial frequency (Sachs, Nachmias, & Robson, 1971). Furthermore, detection thresholds of spectrally complex stimuli such as bars, square wave gratings, rectangular gratings, and sawtooth gratings were shown to be predictable from the detection thresholds of the constituting sinusoidal components (Campbell, Carpenter, & Levinson, 1969; Campbell & Robson, 1968). From these findings, Robson and others inferred that the visual system initially represents image information as if it performed a spatially localized Fourier analysis. Inspired by the demonstration of form-selective responses by neurons in primary visual cortex (Hubel, 1959; Hubel & Wiesel, 1962, 1968), they proposed a model for human pattern detection in which visual choice behavior is directly linked to the underlying single cell physiology. Specifically, neurons in visual cortex were suggested to constitute functionally separate mechanisms ("channels") that act as an array of independent and linear filters with rectified outputs. Perceptual choices in psychophysical tasks that involve visual patterns at near-threshold contrasts were suggested to result from evaluating the multichannel output in a simple way, for instance, by comparing the output of the maximally responsive channel (Graham, 1989; Pelli, 1985) or the Minkowski pooled responses of all channels[2] (Stromeyer & Klein, 1975) to a decision criterion.

The successful application of linear systems analysis to behavioral pattern detection provided a unifying mathematical framework to describe visual information, neural computation, and psychophysical sensitivity and prompted major research efforts in vision science. Before long, electrophysiologists applied the power of linear systems analysis techniques to quantitatively study the computations performed by simple and complex cells in primary visual cortex (Movshon, Thompson, & Tolhurst, 1978a, 1978b) and motion direction selective neurons in the middle temporal visual area (MT; Movshon, Adelson, Gizzi, & Newsome, 1985). At the behavioral level, hundreds of near-threshold studies reported results that were qualitatively and quantitatively well described by linear channel models adhering to the basic architecture sketched in Figure 1 (for an excellent review on this vast psychophysical literature, see Graham, 1989). Theoretical work further demonstrated that a spatially localized Fourier analysis provides a sparse code for natural images (Olshausen & Field, 1996; Rehn & Sommer, 2007). In all of these fields, current models and system identification techniques have added several layers of complexity to the original linear channels proposal. Nevertheless, many concepts at the heart of modern vision science have their roots in the pattern detection literature.

Limitations of the Linear Channels Model

Despite its ability to describe a wide range of near-threshold results, the characterization of encoding and decoding employed by the linear channels model is fundamentally at odds with recent insights and models in visual neuroscience. First, spatial-

---

[2] A Minkowski metric has the form

$$R = \left[\sum_i |r_i|^\beta \right]^{1/\beta},$$

where r_i are individual channel responses and β is the pooling exponent.
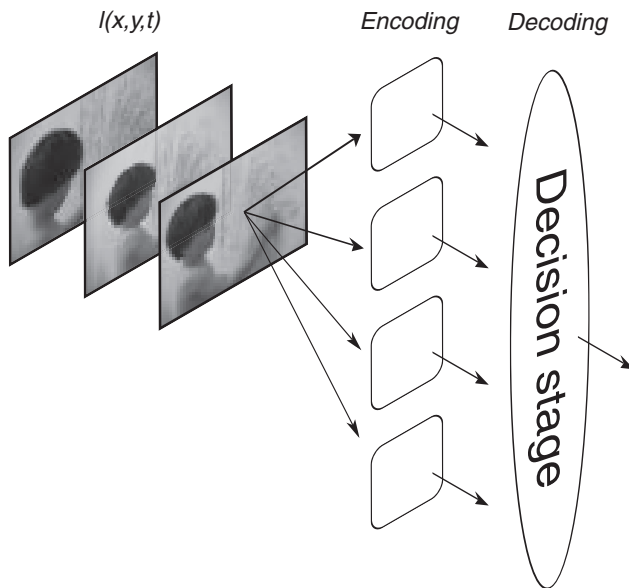
I(x,y,t)   Encoding   Decoding

Figure 1. Cartoon illustration of the linear channels model for threshold pattern vision. In the late 1960s and early 1970s, psychophysicists popularized a quantitative information processing approach to study pattern vision. This framework assumes that the visual system initially analyses the retinal input using parallel pathways tuned to a variety of visual dimensions, left unspecified in this cartoon (for in-depth discussion of these dimensions, see Adelson & Bergen, 1991; Graham, 1989). Visual encoding is carried out by psychophysical channels, assumed to operate linearly and independent of each other and originating in the neural computations performed in visual cortex. In behavioral experiments involving near-threshold stimuli, the decoding stage that connects channel activity to perceptual judgments is typically formalized by a simple task-blind read-out rule such as a maximum-output operator or Minkowski pooling.

frequency-tuned neurons in primary visual cortex are neither linear nor independent encoders due to squaring and gain control mechanisms (Carandini, Heeger, & Movshon, 1997; Heeger, 1992a, 1992b). Squaring produces nonlinearly increasing cell responses as stimulus contrast increases from zero, whereas gain control—achieved via nonspecific suppression—renders neural responses dependent on the activity of neurons tuned to a broad range of stimulus characteristics (see Figure 2). Second, the winner-takes-all decoding strategy that is often employed in the linear channels model is difficult to reconcile with a basic operating principle of the brain: Information is represented in large numbers of neurons, and likewise it seems as if large numbers are taken into account when making perceptual decisions (Parker & Newsome, 1998; Pouget, Dayan, & Zemel, 2003). Moreover, ample evidence emphasizes that even in simple perceptual discrimination tasks, responses of multiple sensory neurons may be pooled according to their reliability and relevance to the task at hand (Fetsch, Pouget, DeAngelis, & Angelaki, 2012; Law & Gold, 2008; Parker, Krug, & Cumming, 2002; Purushothaman & Bradley, 2005). This suggests that behavioral performance is principally limited by noise correlations between neurons (Averbeck, Latham, & Pouget, 2006; Shadlen, Britten, Newsome, & Movshon, 1996; Zohary, Shadlen, & Newsome, 1994), and pooling inefficiencies (Beck, Ma, Pitkow,

Latham, & Pouget, 2012; Cohen & Newsome, 2009; Palmer, Chen, & Seidemann, 2007; Putzeys, Bethge, Wichmann, Wagemans, & Goris, 2012), but not by the most sensitive neuron available (Barlow, 1972) or a task-blind decoding rule as assumed in the linear channels model.

Over the last three decades, many psychophysical studies have pointed out that the encoding and decoding principles put forward in the linear channels model fail to generalize to suprathreshold vision and perceptual tasks other than detection. Detection performance, for instance, increases nonlinearly with stimulus contrast (Legge, 1978; Nachmias & Sansbury, 1974). Moreover, pattern masking studies have convincingly demonstrated the existence of psychophysical parallels of V1 nonlinear response mechanisms, that is, response expansion (Goris, Wagemans, & Wichmann, 2008; Legge & Foley, 1980) and broadly tuned response suppression (Foley, 1994). In a rather different research domain, behavioral experiments addressing the question how the brain solves inference problems typically find that perceptual decision making often follows the normative Bayesian framework. In a wide variety of tasks, human observers take the reliability of sensory information and prior statistical knowledge about the task to be performed into account when making perceptual judgments (Ernst & Banks, 2002; Girshick, Landy, & Simoncelli, 2011; Mamassian & Landy, 2001; Najemnik & Geisler, 2005; Weiss, Simoncelli, & Adelson, 2002).

The relevance of the linear channels model is thus limited to the threshold situation. For this reason, it has often been (tacitly) assumed that the visual system may operate in a special regime in detection tasks (Graham, 1989, pp. 11–12, 1992, 2011; Wandell, 1995, pp. 195–196). Specifically, near-threshold stimuli may be too weak to drive the early visual system's nonlinearities and activate only a very limited number of V1 neurons. In this scenario, maximum-output decoding of linear and independent channels could be an adequate description of the brain's operation in detection tasks. Although this reasoning provides a justification for the use of linear channels models, it severely limits the potential of near-threshold pattern detection tasks as a model system for everyday vision and sensory information processing in general. Alternatively, the shortcomings outlined in this section may be taken to imply that detection calls for a different explanation than the one provided by the linear channels model. That is the assessment we make here, and it provides the direct motivation for this article. Building on the strengths of the early theoretical framework, we seek to identify a computational theory for pattern detection that overcomes its weaknesses.

Current Observer Models for Pattern Detection

Current observer models for pattern vision have overcome some of the limitations of the linear channels model. This article is concerned with behavioral detection experiments. Because pattern detection is a special case of masking (i.e., it can be considered a masking experiment whereby the mask contrast equals zero), we focus exclusively on models developed in the domain of pattern masking in this section (for a more exhaustive review of current pattern vision models, see Graham, 2011). Masking experiments investigate the visual mechanisms that mediate discrimination of a pattern from a mask whereby both stimuli are typically coincident in space and simultaneous in time. The pattern stimulus usually
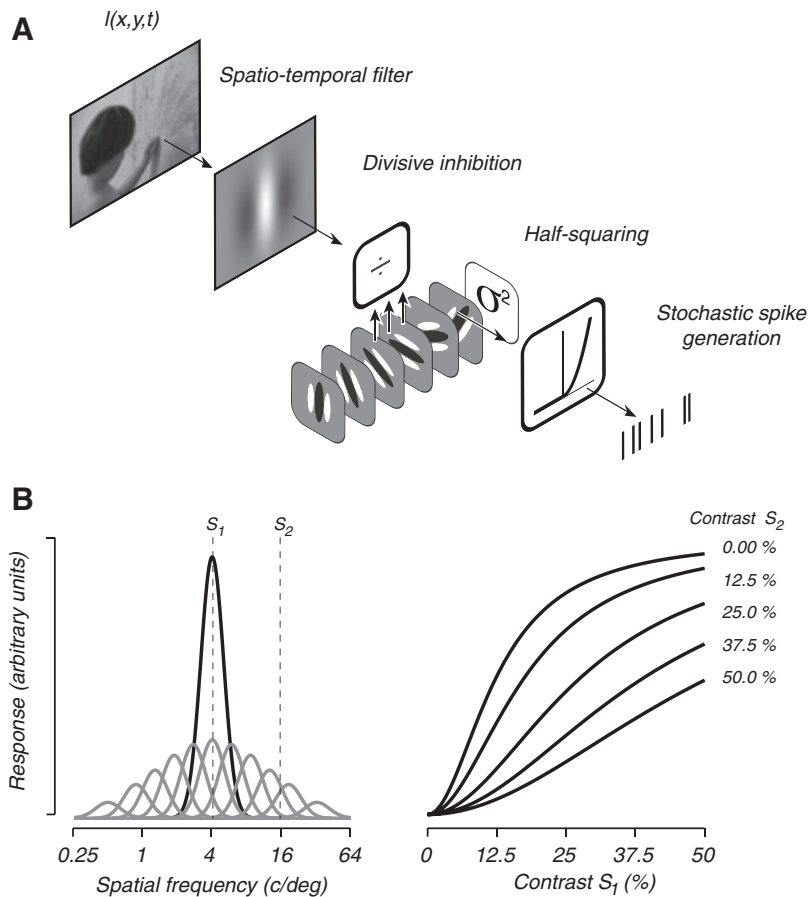
Figure 2. (A) The standard functional model of V1 neurons consists of a linear filtering stage followed by broadband divisive inhibition and response expansion. The variability in neural responses is attributed to a stochastic spike generation process in this model, although this does not accurately reflect the underlying biophysics (action potential generation is a highly reliable event; the noise is in the synaptic inputs). The linear filter shown is a log-Gabor receptive field. (B) V1 neurons have several nonlinear response properties that are well described by this model. The left figure illustrates spatial frequency tuning of the linear filters that excite (black) and suppress (gray) an idealized V1 neuron. Stimulus S_1 drives the neuron vigorously, and S_2 fails to elicit a response. The plot on the right illustrates responses to a family of stimuli composed by weighted summation of S_1 and S_2. The neuron's response to S_1 is modified by the presence of S_2 in a contrast-dependent manner. c/deg = cycles/degree.

consists of a sine wave grating and is commonly referred to as the signal. The mask can vary in a number of ways. Depending on the research question addressed, it may be identical to the signal except for its contrast (e.g., Bird, Henning, & Wichmann, 2002; Nachmias & Sansbury, 1974), differ in multiple aspects (Foley, 1994; Legge & Foley, 1980), or be spectrally complex and stochastic (Baker & Meese, 2012; Henning & Wichmann, 2007; Lu & Dosher, 2008; Nagaraja, 1964; Putzeys et al., 2012). The latter case is sometimes called detection-in-noise. The effect of one stimulus on the detectability of another is often disruptive, although low contrast maskers can increase stimulus detectability (Blackwell, 1998; Goris, Zaenen, & Wagemans, 2008; Nachmias & Sansbury, 1974).

Most observer models are concerned with both masking effects. Following a long tradition, they commonly include a perceptual template (i.e., a linear filter selective for the signal stimulus), a

nonlinear transducer, a gain control process, internal noise, and a decision mechanism (Dosher & Lu, 1998; Foley, 1994; Goris, Zaenen, & Wagemans, 2008; Legge & Foley, 1980). For our purposes, it suffices to consider one example model in more detail. We opt for the perceptual template model (PTM), introduced by Dosher and Lu (1998). It is an extension of many other models and can be applied to both stochastic and nonstochastic masks (Dao, Lu, & Dosher, 2006; Klein & Levi, 2009). For a recent review of the PTM and related models, see Lu and Dosher (2008).

In the PTM, stimuli are processed in two pathways (see Figure 3). The signal pathway has a perceptual template that responds to a signal stimulus but not to stimuli that are very different from the signal. The exact characteristics of the template are usually left unspecified, but it is assumed that these may be highly complex and, for instance, match an object or a face (Macke & Wichmann, 2010). The template output is further processed by a nonlinear
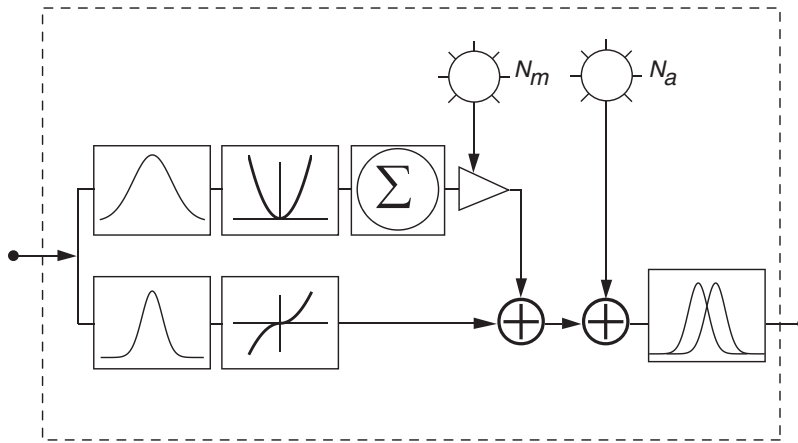
Figure 3. The perceptual template model of a human observer. This schematic model representation is a replica of Figure 15 in Lu & Dosher (2008). The lower pathway is the signal pathway. Visual input is processed by a linear template, the output of which is nonlinearly transduced and subject to the disruptive influence of the gain control pathway. This process can be formalized as a multiplicative noise injection or as a divisive contrast gain control operation. It is followed by an additive internal noise source and a standard one-dimensional signal-detection-theory decision stage (Tanner & Swets, 1954). From "Characterizing Observers Using External Noise and Observer Models: Assessing Internal Representations With External Noise," by Z.-L. Lu and B. A. Dosher, 2008, *Psychological Review, 115,* p. 60. Copyright 2008 by the American Psychological Association.

transducer function. Stimulus gain is modified through the gain control pathway, which also has a template and a nonlinearity that may or may not differ from the signal pathway template and nonlinearity. Together, these pathways give rise to a one-dimensional internal stimulus representation that is passed to the decision stage. Observer sensitivity is limited by internal noise that is produced by an additive and a multiplicative noise source, whereby the latter's magnitude is determined by the total amount of contrast energy in the input stimulus and hence provides the contrast gain control signal. Overall, the PTM can fully characterize an observer using as little as five free parameters (Lu & Dosher, 2008), although there are several ways to extend the model.

The nonlinear response transduction and contrast gain control signal of the PTM connect the model to the functional mechanisms that underlie nonlinear response properties of cortical cells. Moreover, these nonlinearities enable the PTM to parsimoniously account for several detection and masking phenomena. Specifically, the model can simultaneously explain why detection performance increases nonlinearly with stimulus contrast, why stimulus detectability can benefit from the presence of a low-contrast mask,[3] why discrimination thresholds in many visual tasks follow Weber's law, and why response consistency in double-pass experiments increases with task performance (Klein & Levi, 2009; Lu & Dosher, 2008).

Limitations of the PTM

The original intent of the PTM was to provide a simple approximation for stochastic multichannel models that allows to both quantify general perceptual limitations and characterize the nonlinear properties of an observer's internal stimulus representation (Lu & Dosher, 2008). The model therefore provides a rather abstract description of visual processing whereby encoding and

decoding are intertwined.[4] However, leaving the information processing framework of mechanistic models comes at a cost. First, it implies that the PTM response to a visual stimulus cannot in any simple way be related to the pooled responses of a population of visual neurons. The reason is that the PTM nonlinearities are applied to the output of a single linear template. Following Morgenstern and Elder (2012), we consider this argument in detail for the spatial dimension, but it holds for all tuning dimensions that precede the response nonlinearity. Neurons have highly localized and overlapping receptive fields. Even a small Gabor stimulus drives thousands of V1 neurons (Chen, Geisler, & Seidemann, 2006). Outputs of many neurons thus need to be pooled to construct a template. If these outputs were linear, their sum would be indistinguishable from the template response, provided that recep-

---

[3] The performance-enhancing effect of low-contrast masks has been studied extensively for nonstochastic masks (Foley, 1994; Henning & Wichmann, 2007; Legge & Foley, 1980; Nachmias & Sansbury, 1974), but has received less attention for detection-in-noise (but see Blackwell, 1998; Goris, Zaenen, & Wagemans, 2008). Most threshold-versus-contrast (TvC) plots of the PTM do not show this effect. However, Goris, Zaenen, and Wagemans (2008) demonstrated that an extended PTM successfully fits the effects of weak and strong noise, and Klein and Levi (2009) showed that the stochastic version of the PTM predicts beneficial effects of subthreshold noise in detection. In most PTM applications, TvC plots are generated from the simplified analytic version of the PTM, and in this version of the model this effect does not occur.

[4] The PTM should probably be considered a statistical rather than a mechanistic model, which means that it is well defined, has appealing computational properties, and provides principled ways to extract patterns from behavioral choice data. However, it does not attempt to capture in detail how a psychophysical template arises from neuronal responses. The linear channels model, on the other hand, leans more toward being a mechanistic model, meaning that it assumes a specific mechanism, intended to resemble the underlying physiology, that generates the outcome in the decision process.

tive fields were to tile visual space perfectly. Formally, let s_i be the spatial profile of the linear receptive field of neuron i, s_{template} the spatial profile of a global linear template, x and y two spatial dimensions, and l the stimulus luminance profile:

$$R_{\text{pooled}} = \sum_{i=1}^{n} \iint s_i(x,y)l(x,y)dxdy \qquad (1)$$

cannot be distinguished from

$$R_{\text{template}} = \iint s_{\text{template}}(x,y)l(x,y)dxdy. \qquad (2)$$

However, when we introduce the previously discussed nonlinear response transformations, this equality no longer holds. Consider the example of a response exponent. It is clear that for all $p \neq 1$,

$$R_{\text{pooled}} = \sum_{i=1}^{n} \left(\iint s_i(x,y)l(x,y)dxdy \right)^p \qquad (3)$$

will differ from

$$f(R_{\text{template}}) = \left(\iint s_{\text{template}}(x,y)l(x,y)dxdy \right)^p. \qquad (4)$$

Models that are characterized by a global linear template matching stage can therefore not directly be related to models that linearly combine outputs of nonlinear subunits and hence not to neural population activity. For which aspects of detection does the simplified observer model provide a good approximation to a full model? Morgenstern and Elder (2012) recently asked whether assuming a global linear template compromises the ability of observer models to characterize human choice behavior in a detection-in-noise task when the signal is spatially extended. They compared a global linear template model with a local nonlinear subunit model. The latter one turned out to be a much better predictor of human choice behavior on a trial-by-trial basis (see Figure 4).

A second limitation that originates in the statistical nature of the PTM is that the model is agnostic about the mechanisms that shape the perceptual template or the different factors that contribute to internal noise. It is therefore not clear what to expect when we combine two signals or introduce stimulus uncertainty. Specifically, let θ be the vector that contains the parameter estimates of the PTM that account for detection of S_1 and θ' the parameter vector for detection of S_2. Although the observer's visual processing of S_1 and S_2 is fully characterized, we cannot in a principled way derive a prediction from θ and θ' for a summation experiment, that is, detection of $S_3 = S_1 + S_2$, or for an uncertainty experiment, that is, detection of S_1 and S_2, when $P(S_1) = q$ and $P(S_2) = 1 - q$. It is exactly this type of
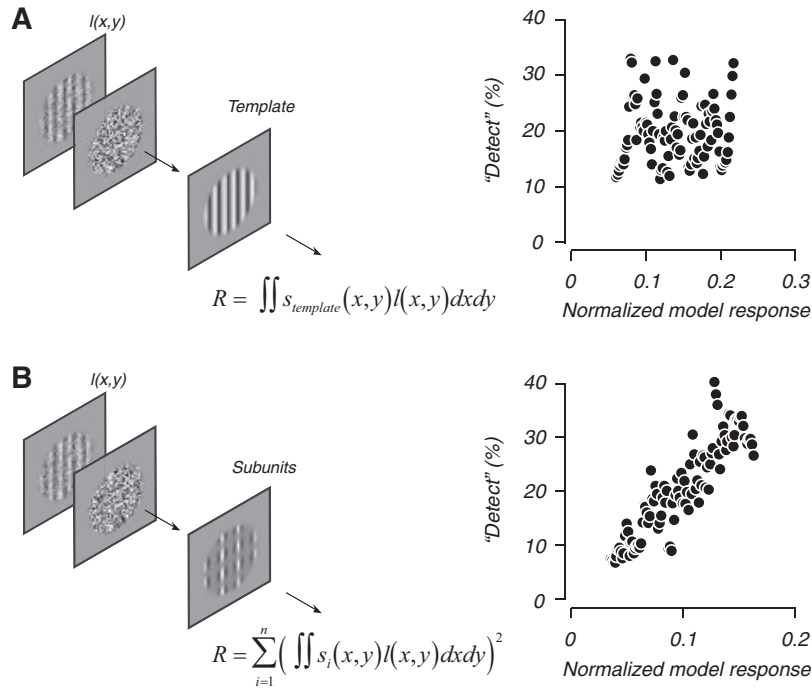

Figure 4. Template matching models fundamentally differ from subunit models when visual encoding is nonlinear. Morgenstern and Elder (2012) examined how well both types of models predict human choice behavior in a detection-in-noise task that involved spatially extended stimuli. The signal-absent trials are of particular interest, as they unambiguously reflect which image characteristics lead human observers to "see" the signal in the noise. (A) Template matching models by nature assume that spatial integration can be described as a process that precedes nonlinear response transformations. However, responses of a matched linear filter are a very poor predictor of human choice behavior (only signal-absent trials are shown). This was also true for a global mechanism built from templates with different phase tuning (not shown). (B) A subunit model in which the outputs of local energy mechanisms are summed turns out to give a much better characterization of perceptual judgments.

connection between experiments we seek to establish in this article. This calls for a multichannel approach, whereby the encoding and decoding stage are separated and made explicit (for simulations and analysis of a multichannel variant of the PTM in the context of perceptual learning, see Dosher & Lu, 1998; Petrov, Dosher, & Lu, 2005).

Desiderata for a Theory of Pattern Detection

Here we seek to formulate a computational theory of pattern detection that combines the strengths of the linear channels model and the PTM and overcomes their weaknesses. On the basis of the arguments presented above, we identify the following desired characteristics of this theory. First, in keeping with the main strength of the linear channels model, the theory should be explicitly connected to the physiological underpinnings of pattern vision. Second, following the PTM, it should be founded upon computational principles that link detection with suprathreshold pattern vision. Third, because pattern detection is simply one of many sensory information processing tasks, the decoding mechanism ought to generalize to perceptual decision making in general. And fourth, in line with the aspirations of the linear channels model and the PTM, it should explain a broad range of behavioral results using a single set of parameters.

With these goals in mind, we have selected a set of classical pattern detection experiments (see Table 1) that seek to answer three fundamental questions: How is detectability of a compound stimulus related to detectability of its components? How is stimulus detectability modified by recent sensory experience? And how does detectability depend on stimulus probability? These questions have been addressed in summation, adaptation, and uncertainty experiments, respectively, and will form the test bed for the model that we outline in the following section and examine in the rest of this article.

A Neural Population Model for Pattern Detection

In the preceding sections, we motivated the need for a unified theory of pattern detection and argued that several current observer models of pattern masking do not possess the full set of characteristics needed to be considered viable candidates. In the remainder of this article, we describe and evaluate a theory that we refer to as the neural population model (NPM). In this model, visual pattern detection is formalized as an information processing problem; that is, the multidimensional stimulus encoding and dimensionality-reducing decoding stage are separated and made explicit. Visual processing is performed by units whose basic stimulus selectivity is determined by a linear filtering operation. The filter responses are subject to nonlinear transduction (squaring) and divisive inhibition and are corrupted by noise modeled to reflect neural noise. The sensory population response is read out by a linear decoder that aims to maximize task performance without having detailed knowledge about the structure of the full response distribution. A schematic of the model is shown in Figure 5.

Models that bear similarity to ours have been applied before in the domains of pattern masking (Chirimuuta & Tolhurst, 2005; Geisler & Albrecht, 1997; Goris, Wichmann, & Henning, 2009; Itti, Koch, & Braun, 2000; Sanborn & Dayan, 2011), perceptual learning (Petrov et al., 2005), and visual attention (Lee, Itti, Koch,
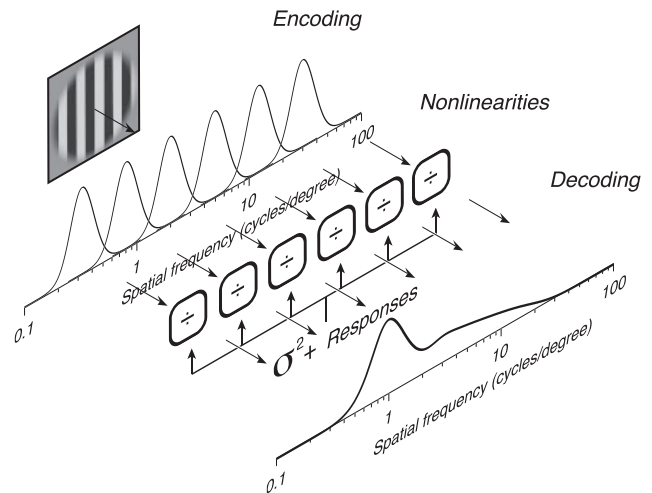

Figure 5. The neural population model of pattern detection. Visual information is analyzed by a population of spatial frequency tuned units. As in the standard functional model of V1 neurons, linear filtering is followed by divisive inhibition and response squaring (not shown). Unit responses are corrupted by weakly correlated multiplicative noise (not shown) and read out by a linear maximum-likelihood decoder that makes some simplifying assumptions about the structure of the full population response distribution.

& Braun, 1999). In all these models, visual information is initially processed by a population of linear–nonlinear cascades tuned to the stimulus dimensions of interest. However, the models do differ in the details of the encoding stage, in the assumed characteristics of internal noise, and in the postulated mechanisms for decision making. Nevertheless, the similarity of approach does suggest that it may be possible to formulate a single multichannel model capable of simultaneously explaining important aspects of visual attention, perceptual learning, masking, and, as we shall see, pattern detection.

Mathematical Properties of the Neural Population Model

Stimulus Encoding

In the NPM, visual input is encoded by a population of spatial-frequency-tuned units. To approximate the cortical input signal provided by thalamic neurons in the lateral geniculate nucleus (LGN), the first processing stage in the model consists of applying a linear bandpass filter M to the input. M encompasses the effects of the eye's modulation transfer function as well as the spectral filtering that takes place in retinal ganglion cells and lateral geniculate nucleus cells. Let θ be spatial frequency, then contrast attenuation is given by

$$M(\theta) = \theta^\alpha e^{\beta\sqrt{\theta}}. \qquad (5)$$

Parameters α and β control the shape of the bandpass filter and were determined through fitting a large set of behavioral data (see Table 2). For convenience, the output of M is scaled between 0 and 1.

The second processing stage in the model is designed to mimic the neural computation taking place in V1. As has become standard, this step is characterized by a linear–nonlinear cascade model. The units' basic selectivity is thus determined by a linear receptive field, chosen to exhibit a log-Gaussian-shaped spatial-frequency-tuning function f:

$$f(\theta) = e^{-\frac{(\ln\theta - \ln\theta_c)^2}{\phi}}, \qquad (6)$$

where θ is spatial frequency, θ_c the unit's preferred spatial frequency, and ϕ the parameter that controls tuning width. The response of f is scaled between 0 and 1. The units' preferred spatial frequencies $\theta_c = \{\theta_{c1}, \ldots, \theta_{cn}\}$ are evenly distributed on a logarithmic scale over the range [0.1, 66] cycles degree^{-1}. Parameter ϕ was determined through fitting (see Table 2).

Let L_i refer to the response of the ith unit's linear receptive field. The mean unit response rate R_i is a nonlinear transformation of L_i reflecting the effects of nonlinear transduction and broadly tuned divisive inhibition:

$$R_i = r_0 + r_{\max}\left(\frac{L_i}{\sqrt{\sigma^2 + \sum_{j=1}^{m} z_j L_j^2}}\right)^n. \qquad (7)$$

Here r_0 is the ith neuron's mean spontaneous discharge rate, $r_{\max}$ its maximal firing rate, n its response exponent, and σ its semi-saturation contrast. The normalization signal is computed by taking the weighted sum of all squared filter responses, whereby the weights $z = \{z_1, \ldots, z_n\}$ follow a log-Gaussian profile centered at θ_c. We set r_0 to 5 impulses per second, assumed perfect squaring for the nonlinearity (i.e., $n = 2$), and used a 4-octave-wide profile for the weights z. The normalization weights were scaled such that a preferred stimulus produces a normalization response of 1. Parameters σ and $r_{\max}$ were determined trough fitting (see Table 2). Response rates are converted to spike counts by multiplying R_i with stimulus duration t.

The encoding stage in our model does not include a description of the units' orientation selectivity, temporal selectivity, or spatial receptive field envelope. The motivation for this simplification is twofold. First, all experiments discussed in this article solely involve manipulations of spatial frequency. Second, detection performance in our model is primarily driven by appropriately tuned units and not affected by inclusion of uninformative units.

Sensory Noise

Visual neurons produce highly fluctuating responses to repeated presentations of the same stimulus (Churchland et al., 2010; Tolhurst, Movshon, & Dean, 1983). For all stimuli, response variance is assumed to be Gaussian and level dependent, that is,

$$\text{VAR}(tR) = ktR, \qquad (8)$$

where k is a proportionality constant (the Fano factor) set to 1.5. We further assume neurons to have weakly correlated trial-to-trial variability in their spike counts, in line with numerous electrophysiological observations (e.g., Kohn & Smith, 2005; see Cohen & Kohn, 2011, for a recent review), although these findings have recently been contested (Ecker et al., 2010). Such correlated response fluctuations are thought to reflect the functional connectivity of neural circuits. For this reason, the strength of the pairwise correlations in our model is maximal for neurons preferring the same spatial frequency and decreases monotonically as the preferred frequencies become more different (see Figure 6). We based the values of the noise correlations on reported estimates in the literature (see Table 2; Cohen & Kohn, 2011).

Stimulus Decoding

In detection tasks, the visual system needs to disentangle subtle stimulus-driven activity of sensory neurons from spontaneous neural firing. This is a difficult problem because neu-

Table 2
Parameters Used in the Neural Population Model

Symbol	Value	Comments
α	**1.91**	Controls the shape of the modulation transfer function M
β	**−2.27**	Controls the shape of the modulation transfer function M
ϕ	**1.01**	Full bandwidth at half height for the units' linear receptive field, expressed in octaves
$r_{\max}$	**194.9**	Maximal firing rate, expressed in impulses per second
σ	**0.015**	Stimulus-independent response suppression
γ	**8.14**	Controls the strength of neural adaptation for adequately stimulated neurons
δ	**3.22**	Controls the strength of neural adaptation for weakly stimulated neurons
ϵ	**59.9**	Maximal duration of response integration during adaptation, expressed in seconds
r_0	5.0	Spontaneous firing rate, expressed in impulses per second
n	2.0	Response exponent, chosen to produce energy mechanisms
k	1.5	Fano factor
θ_c	[0.1, 66]	Minimal and maximal preferred spatial frequency, expressed in cycles per degree
ρ_{ij}	[0.05, 0.15]	Minimal and maximal noise correlation between sensory units

Note. Values of the free parameters, obtained through fitting the behavioral data, are shown in bold.
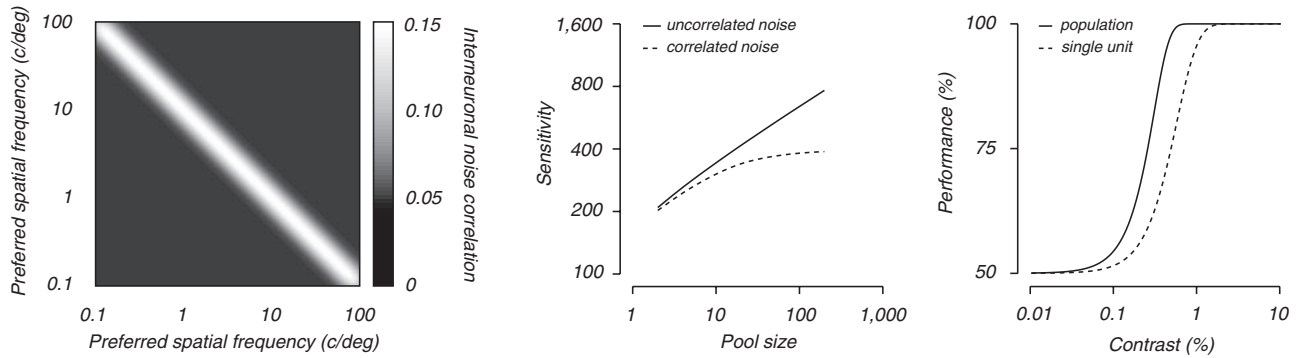
Figure 6. The performance limiting effect of correlated sensory noise on detection. The left panel illustrates the interneuronal correlation coefficient matrix implemented in the neural population model. Correlation strengths vary between 0.05 and 0.15, are maximal for units with the same preferred spatial frequency, and decrease with a 1-octave-wide log-Gaussian profile for pairs of neurons with different preferred frequencies. The middle panel shows how sensitivity (i.e., the inverse of the 75% correct detection threshold) evolves as a function of pool size with and without correlated noise included. All neurons in the pool were tuned to the signal in this simulation. When their responses are correlated, detection performance saturates at a pool size of approximately 50 units. This saturation reflects the fact that pooling neural responses averages out the independent noise within the pool but leaves the correlated component unaffected. Detection thresholds of the saturated pool correspond to the performance level that is reached by optimally combining seven independent units, in line with the measurements in awake behaving monkey of Palmer et al. (2007). The right panel shows a neurometric function for a single unit and the psychometric function of the saturated population, consisting of 200 units. The psychometric function has a slope that is steeper than the neurometric function, consistent with the observations of Palmer et al. c/deg = cycles/degree.

rons vary in their level of maintained discharge, produce highly fluctuating responses to repeated presentations of the same stimulus, and share some of their noise. Typically, only a small fraction of cells is selective for the stimulus. All these characteristics are present in our model. The optimal inference strategy to maximize performance in a two-alternative forced-choice (2AFC) detection task consists of choosing the interval most likely to contain the signal given both population responses (i.e., maximum likelihood decoding). Computationally, this can be achieved by linearly combining the spike counts using a set of weights $\mathbf{w} = \{w_0, \ldots, w_n\}$ determined by the product of each unit's mean increase in responsiveness when the signal is present and the reliability of this increase (Green & Swets, 1966). Let $r(s)$ and $r(b)$ refer to the *i*th neuron's response to the stimulus and a blank, then the optimal linear weight is

$$w_i = \frac{E[r(s) - r(b)]}{\text{VAR}(r(s) - r(b))}. \quad (9)$$

Choosing the stimulus alternative that elicited the larger weighted sum

$$D = \sum w_i t R_i \quad (10)$$

maximizes detection performance, provided that responses are normally distributed and statistically independent. This latter assumption is of course violated in our model. Importantly, even very weak correlations between individual neurons yield highly correlated population activity.

Because our decoder does not take the structure of the response covariance matrix into account, correlated neural noise effectively provides an upper limit to signal detectability in the

NPM (see Figure 6). Although clearly suboptimal in a statistical sense, this may be regarded as a realistic model assumption: Estimating the covariance matrix for large populations of neurons simply is not feasible. An argument in support of this notion is provided by the fact that the relation between neurometric and psychometric functions in the NPM closely resembles measurements in macaque monkeys performing a detection task. The psychometric function is characterized by a threshold that is slightly lower and a slope that is slightly higher than the corresponding values of the neurometric function (Palmer et al., 2007).

In sum, the assumed decoder is suboptimal because it does not attempt to extract information from the correlated sensory noise (Chen et al., 2006; Graf, Kohn, Jazayeri, & Movshon, 2011). However, this omission has the benefit of computational simplicity and successfully captures the empirically estimated relation between sensory unit sensitivity and behavioral performance in contrast detection tasks (see Figure 6).

Detection Performance

Model performance in a 2AFC detection task depends on the decision statistic D, that is, a weighted sum of spike counts. We assumed no temporal uncertainty and a maximal integration window of 100 ms (Gorea & Tyler, 1986). Neither of these assumptions is critical for the model behavior: Including temporal uncertainty or decreasing the temporal integration window would simply increase the contribution of the spontaneous activity to the decision statistic and hence increase detection thresholds. For any given stimulus, D will be distributed normally if responses are combined linearly. Given that this con-

dition is satisfied in the NPM, stimulus detectability is fully determined by the mean and variance of D. Let $D(s)$ refer to the signal present interval. With Equation 9 to infer the weight attributed to each unit,

$$E[D(s)] = \sum_{i=1}^{N} w_i r_i(s) \qquad (11)$$

and

$$\text{VAR}(D(s)) = \sum_{i=1}^{N} w_i^2 \text{VAR}(r_i(s))$$
$$+ \sum_{i=1}^{N} \sum_{j \neq i}^{N} \rho_{ij} \sqrt{w_i^2 \text{VAR}(r_i(s)) w_j^2 \text{VAR}(r_j(s))}, \qquad (12)$$

where N is the number of units in the pool and ρ_{ij} the correlation coefficient between the ith and the jth unit. Mean and variance of D for the signal absent interval (the "blank"), $D(b)$, are derived by replacing $r(s)$ with $r(b)$ in Equations 11 and 12. Detection performance, expressed as d'—i.e., the signal-to-noise ratio of an ideal observer—can be computed directly from these statistics:

$$d' = \frac{E[D(s)] - E[D(b)]}{\sqrt{\dfrac{\text{VAR}(D(s)) + \text{VAR}(D(b))}{2}}}. \qquad (13)$$

Model Fitting

The NPM was fitted to several data sets from different experiments simultaneously with a single set of parameters (see Tables 1 and 2). To ensure that the model was operating in the saturated sensitivity regime, we included 200 sensory units with varying stimulus selectivities. The model makes use of eight free parameters in total. Two of these (α and β) control the global spatial frequency filtering that takes place in the first processing stage; one (ϕ) the bandwidth of the units' linear receptive field; one (σ) the stimulus-independent suppression; one ($r_{\max}$) the maximal unit response rate; and three (γ, δ, and ϵ) adaptation-induced response reduction.

We minimized root-mean-square error of the logarithmically transformed data and model predictions, $\text{RMSE}_{\ln}$, as given by

$$\text{RMSE}_{\ln} = \sqrt{\frac{1}{N} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left(\ln \frac{\hat{\psi}_{ji}}{\psi_{ji}} \right)^2}, \qquad (14)$$

where j is the experiment index, n_j the number of data points in experiment j, $N = \sum n_j$, ψ_{ji} the behavioral detectability estimate, and $\hat{\psi}_{ji}$ the corresponding model prediction. We express fit error on a logarithmic scale because detection thresholds vary over several log units. The reported RMSE values are computed by taking the exponential of $\text{RMSE}_{\ln}$ and express the distance between model prediction and data as a fraction of the threshold contrast. Finally, detectability ψ is expressed as threshold contrast for sine wave detection and as threshold ratio for all other experiments. Both statistics reflect the same currency, that is, relative pattern visibility.

Empirical Test of the Neural Population Model for Pattern Detection

In this section, we evaluate the theoretical predictions of the NPM against empirical data. All model behavior described originates from a single set of parameters (see Table 2). The free parameters were estimated by fitting the model to a subset of the empirical data (see Table 1) and kept constant for all simulations described in this article. We consider several summation, adaptation, and uncertainty experiments that have investigated manipulations of the spatial frequency dimension. The empirical phenomena we discuss are all well known, though some findings are disputed. For a more extensive review of these empirical findings and their interpretation in the framework of linear channels models, see Graham (1989).

First, consider detection of sinusoidal patterns, which is treated as baseline. Detection thresholds vary a great deal with stimulus spatial frequency. Some part of this variation is caused by the optics of the human eye, although the exact shape of the *contrast sensitivity function* depends on many more factors that are not of direct interest here such as mean luminance (Campbell & Robson, 1968), temporal stimulus profile (Henning, 1988), stimulus size (Luntinen, Rovamo, & Naesaenen, 1995), and the statistics of the visual environment (Bex, Solomon & Dakin, 2009). To calibrate the NPM's absolute contrast sensitivity, we included one well-known example of contrast sensitivity measurements in our fit data set (Campbell & Robson, 1968). As can be seen in Figure 7, the contrast sensitivity function of the NPM—that is, the inverse of the detection threshold versus spatial frequency—is in good agreement with these data (RMSE = 14% of the threshold contrast). Throughout the rest of this article, experimental measurements and model predictions will often be expressed relative to baseline sensitivity.

Summation Experiments 1: Pattern Detectability and Homogeneity

Summation experiments investigate how detectability of a compound stimulus is related to the detectability of its components. There are several variants that have a somewhat different theoretical motivation. The interest in summation properties of the human visual system originates in linear systems analysis. Linear systems have the appealing characteristic that responses to any input stimulus can be straightforwardly predicted from a limited number of measurements. The basic tests of linearity are homogeneity and additivity, which together constitute superposition.

A system is homogeneous if its output scales accordingly with the input. In case of the visual mechanisms mediating pattern detection, homogeneity would hold if pattern detectability (expressed in d') would double if we were to double the signal contrast. This is arguably the simplest type of summation experiment. Empirical tests of this nature have revealed two fundamental characteristics of pattern detection: First, if we make abstraction of absolute stimulus sensitivity and normalize the signal contrast by the stimulus-specific detection threshold, d' versus c' functions are remarkably invariant for sine wave gratings (here c' is normalized contrast). When plotted on
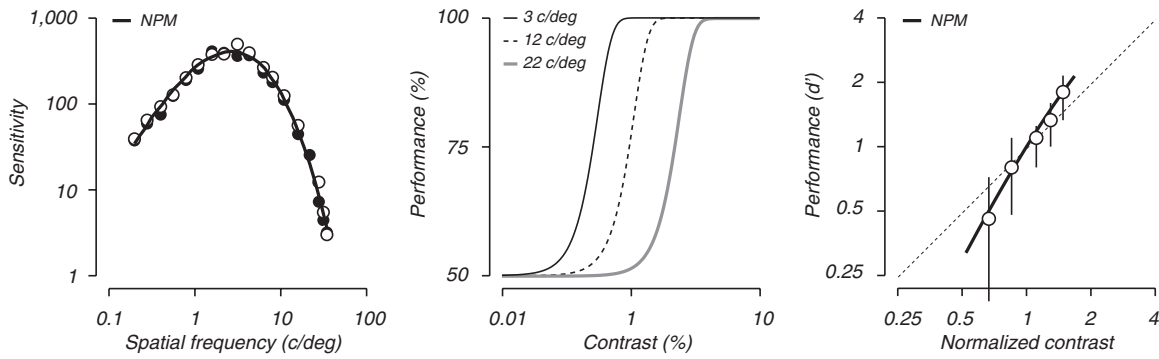
Figure 7. Detection of sinusoidal gratings of varying spatial frequency and contrast. The left panel illustrates the fit of the neural population model (NPM) to the contrast sensitivity function measurements of Campbell and Robson (1968). Sensitivity, that is, the inverse of the 75% correct threshold in a two-alternative forced-choice detection task, is plotted against spatial frequency for observers FWC (black) and JGR (white). The thick black line shows the model prediction. The middle panel shows psychometric functions of the NPM for three spatial frequencies. Consistent with human data (not shown), the slope of the psychometric function does not depend on absolute visibility. The right panel shows a psychometric function of the NPM and a human observer in a different coordinate system—observer WWL in Legge (1978, Figure 7); data replotted from Pelli (1985). Performance is now expressed in d', and contrast in c'. Error bars denote the 95% confidence interval. c/deg = cycles/degree.

semilogarithmic coordinates, the slope of the psychometric function for detection is thus fairly independent of factors such as spatial frequency (Wichmann, 1999, Figures 4–1, 4–4). Second, near threshold, detectability rises faster than linear. Both observations are well captured by the NPM (see Figure 7).

In the NPM, homogeneity for pattern detection does not hold because stimulus contrast is encoded nonlinearly, as is the case in observer models like the PTM (Dosher & Lu, 1998). Note that this is not the only possible interpretation for this failure of linearity. An alternative option that has been studied in detail by Pelli (1985) and Eckstein, Ahumada, and Watson (1997) is the combination of noisy linear encoding mechanisms and a maximum-output decoder. This observer model, known as the uncertainty model, produces similar behavior when the encoding pool is dominated by task-irrelevant units.

Summation Experiments 2: Pattern Detectability and Additivity

A system is additive if its response to the sum of two input stimuli is equal to the sum of each of its responses to these stimuli presented separately. Graham and Nachmias (1971) tested additivity of pattern detection for a large number of stimuli composed of two disparate spatial frequencies, f and $3f$ cycles/degree. Figure 8 illustrates how a high-contrast version of their compound stimulus is encoded in the NPM. Because the fitted bandwidth of the sensory units is 1 octave at half height, the compound activates two subpopulations of units, each tuned to one component frequency. The broadly tuned gain control mechanism suppresses the sensory responses relative to the conditions in which the components are presented separately. The overall population response elicited by the compound thus deviates from the linear expectation; that is, divisive inhibition renders population activity subadditive. Note that this effect is contrast dependent: At low contrasts, the suppression is smaller than at high contrasts (in this case, the denom-

inator in Equation 7 is dominated by the stimulus-independent term σ; also see Figure 2B).

At the behavioral level, the NPM predicts that stimulus detectability is hardly affected by the presence of the second low-
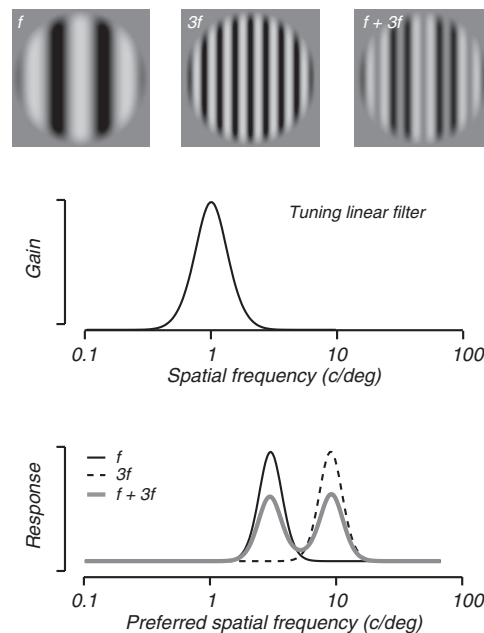

Figure 8. Encoding of stimuli composed of two disparate spatial frequencies in the neural population model (NPM). The upper row illustrates example component stimuli and the resulting compound grating introduced by Graham and Nachmias (1971). To ease illustration, all stimuli are shown at high contrasts. Spatial frequency tuning of the linear filter of the NPM units is shown in the middle panel. The average population response to each of these high-contrast stimuli is plotted in the bottom panel. c/deg = cycles/degree.

contrast grating, which describes the behavioral findings satisfyingly (see Figure 9A; RMSE = 7% of the relative threshold contrast). Detection performance is not hindered by the relative weakening of response strengths because the maximum-likelihood decoder integrates all available information and thus uses the responses to both components to detect the compound signal. As is evident from Equation 9, altering the statistics of the visual input changes the read-out weights employed by the decoder. Crucially, this task-specific decoding originates in a deterministic mechanism that is based on statistical considerations, not on fitting human behavior.

Graham and Nachmias (1971) attributed the behavioral subadditivity to independently operating encoding mechanisms whose output is related to behavioral thresholds via maximum operators. To test over which range of frequencies both components of a complex grating are detected independently, Sachs et al. (1971) measured psychometric functions for a large number of complex gratings with varying spectral separation between the components. Some example predictions for their experiment generated by the NPM are shown in Figure 9C. Sachs et al. found that the independence hypothesis could only be consistently rejected when the spectral separation was smaller than 0.33 octaves. For separations exceeding 0.75 octaves, the improvements in detection performance brought about by addition of the second component were too small to
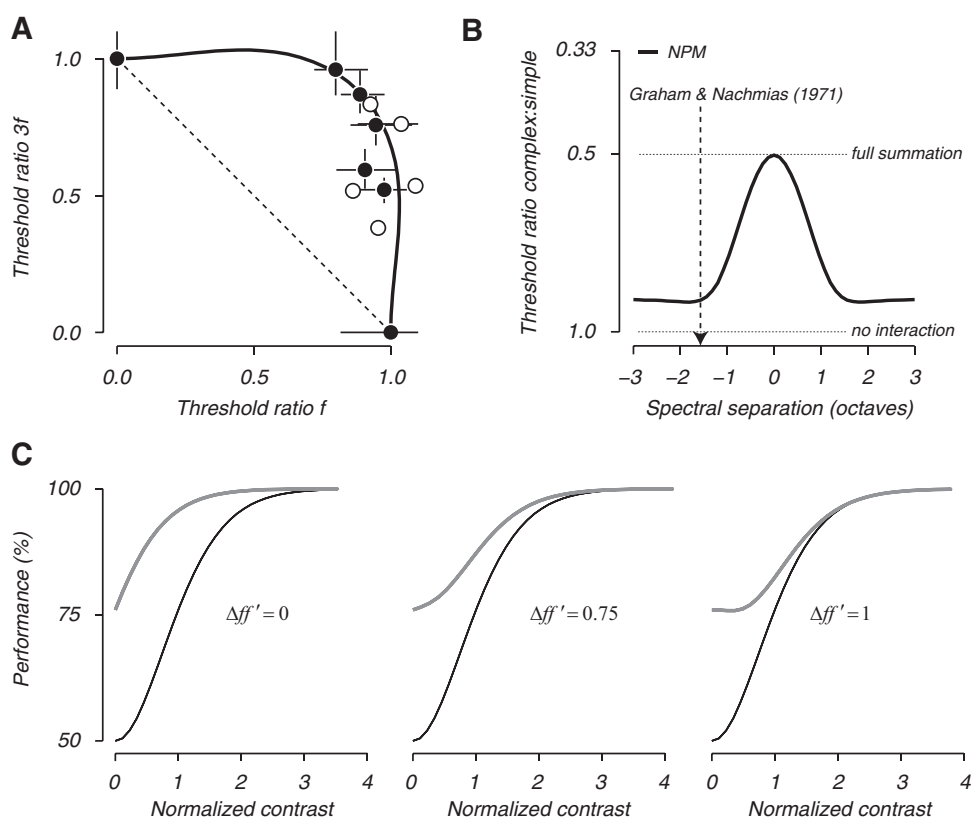

Figure 9. Detection of stimuli composed of two frequencies. (A) The detection threshold of several compound gratings is plotted relative to the threshold for each component measured in isolation. Black symbols are used for observer AG, white for observer RG. The thick black line shows the fit of the neural population model (NPM). The dotted black line illustrates perfect summation. We replicated the experiment performed by Graham and Nachmias (1971) using a slightly different methodology that allowed us to estimate the 68% confidence intervals on the threshold ratios (see Appendix). (B) Threshold ratio of a complex grating and one of its components is plotted as a function of spectral separation of the component gratings for the NPM. Note that the entire region between "no interaction" and "full summation" indicates subadditivity (data outside this region indicate superadditivity and masking, respectively). (C) The design of the experiment performed by Sachs et al. (1971). Simple and complex psychometric functions generated by the NPM are shown in black and gray. The complex psychometric functions illustrate how detection of a compound stimulus improves with the contrast of component f when f' is at its own, independently determined contrast detection threshold. The complex function depends on the spectral separation of the component gratings, $\Delta ff'$, expressed in octaves.
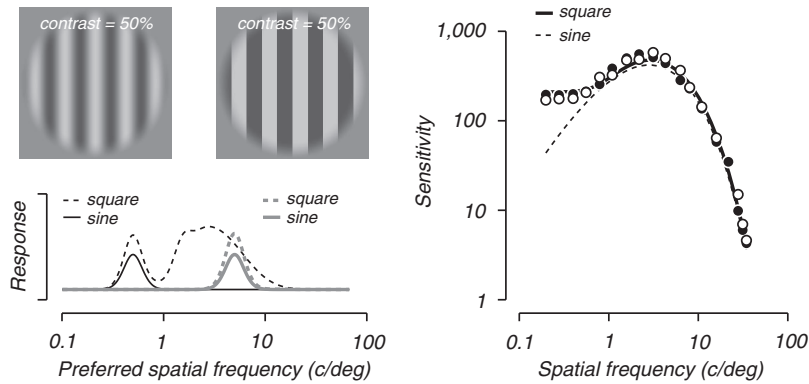
Figure 10. Detection of square wave gratings. The lower left panel illustrates the average population response to threshold contrast sine and square wave gratings of low and high frequency, respectively. The high-frequency stimuli drive the same units, but the different power of their fundamental harmonic leads the square wave to be a more effective stimulus. The low-frequency square wave, on the other hand, drives many more units than its sinusoidal counterpart. The broadband nature of this response is exploited by the maximum-likelihood decoder and results in superior square wave detectability, shown in the right panel. Square wave sensitivity is plotted against spatial frequency for observers FWC (black) and JGR (white). The thick black line shows the fit of the neural population model (NPM). The dotted line plots the NPM's sinusoidal contrast sensitivity function from Figure 7 and serves as reference. c/deg = cycles/degree.

reject the independence hypothesis.[5] The full set of predictions generated by the NPM is summarized in Figure 9B. Threshold ratio is plotted as a function of spectral separation. A ratio of 0.5 corresponds to perfect summation, a ratio of 1 to unaffected detection performance. Consistent with the findings of Sachs et al., maximal summation only occurs at very small spectral separations. The performance improvement falls off rapidly as both components get more separated in the frequency domain and saturates at a level previously attributed to probability summation. Qualitatively, the NPM is thus in good agreement with the data of Sachs et al.

Summation Experiments 3: Pattern Detectability and Superposition

Homogeneity and additivity together constitute the principle of superposition. For a linear system, responses in a superposition experiment can be predicted from the responses to a set of basis functions. As we have concluded in the preceding sections, neither additivity nor homogeneity holds for the human visual system or the NPM. However, because the nonlinearities of the NPM are fully characterized, we can predict the outcome of a superposition experiment. An illustrious historical example is given by Campbell and Robson's (1968) square wave detectability measurements. A square wave grating of fundamental spatial frequency θ and spatial contrast c can be considered as the sum of an infinite series of sine wave components with frequencies equal to the odd harmonics of θ. The contrast of the fundamental component equals $4/\pi \times c$ and decreases gradually for the higher components, that is, $c_h = 4c/\pi h$, where h indicates the odd harmonic.

Figure 10 illustrates encoding of a square wave stimulus in the NPM. Because of the bandpass shape of the contrast sensitivity function in Campbell and Robson's (1968) experiments (see Figure 7), a natural distinction exists between square wave gratings of low and high fundamental frequency. For the latter case, only the

fundamental harmonic is relevant. The NPM's detection thresholds for these square wave gratings are $4/\pi$ times lower than detection thresholds for their sine wave counterparts, which simply reflects the increased power of the first harmonic in the Fourier series. Low-frequency square waves, on the other hand, are effectively broadband and activate many units. This global increase in population activity is exploited by the maximum-likelihood decoder, leading to a dramatic improvement in detectability relative to a narrow-band sine wave grating. As can be seen in Figure 10, square wave sensitivity is correctly predicted by the NPM across the entire range of spatial frequencies (RMSE = 11% of the relative threshold contrast).

Stromeyer and Klein (1975) performed a superposition experiment in which they compared detectability of amplitude-modulated (AM) and frequency-modulated (FM) sinusoidal gratings near the peak of the contrast sensitivity function. Examples of these stimuli and their spectra are shown in Figure 11. The Fourier spectrum of the FM grating consists of a central component at frequency f plus components placed symmetrically either side of f. As is the case for the square wave grating, there is a disconnect between the spatial contrast of the FM grating and the amplitude of

---

[5] The independence hypothesis tested by Sachs et al. (1971) was derived from a high-threshold theory model of detection that can be summarized as follows: (1) visual encoding is performed by spatial-frequency tuned channels; (2) the outputs of these channels are statistically independent; (3) the output of each channel is passed through a separate threshold mechanism; (4) an observer detects a stimulus when the critical level of at least one of the threshold mechanisms is exceeded or when a detect-response is generated by an independent guessing mechanism. According to this model, the probability of detecting a complex stimulus equals one minus the product of the probability of not detecting the first component and the probability of not detecting the second component. The model therefore predicts a small performance improvement when both components drive two independent mechanisms, an effect known as "probability summation."
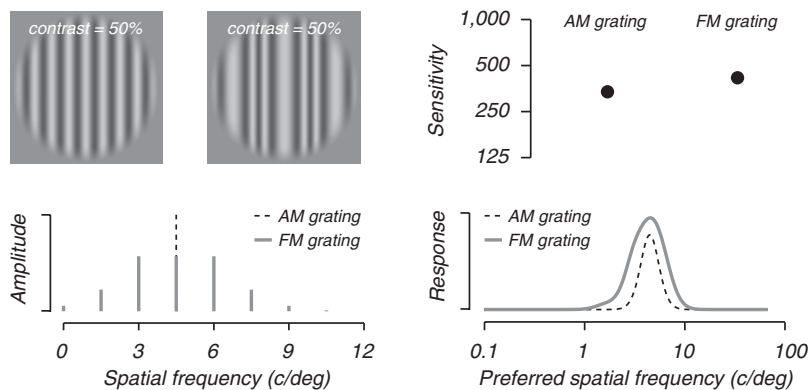
Figure 11. Detectability of amplitude-modulated (AM) and frequency-modulated (FM) gratings. The lower left panel shows the amplitude spectra of the AM and FM gratings used by Stromeyer and Klein (1975). The average population responses to high-contrast versions of both gratings are shown in the lower right panel, and predicted behavioral sensitivity in the upper right panel. c/deg = cycles/degree.

its frequency components. Specifically, a full-contrast FM grating has a maximal amplitude of 0.55, whereas a full-contrast AM grating has a maximal amplitude of 1. Detectability of a square wave grating near the peak of the contrast sensitivity function is well predicted by the power of its fundamental harmonic (see Figure 10). This is not the case for the FM gratings used by Stromeyer and Klein: Despite the considerable amplitude difference, AM gratings and FM gratings are about equally detectable. This result is not inconsistent with predictions of the NPM, shown in Figure 11. In the model, both gratings are encoded differently: The population responses differ (mildly) in width and amplitude. Consequently, the pooling rules employed by the maximum-likelihood decoder will differ somewhat as well (see Equation 9). Detectability, however, does not differ very much: The FM grating yields a detection threshold that is only 19 % smaller than its AM counterpart.

In this section, we have demonstrated that the NPM successfully predicts the outcome of summation experiments designed to test homogeneity, additivity, and superposition of threshold vision. We now turn our attention to adaptation experiments.

Adaptation Experiments 1: A Simple Model for Cortical Adaptation

Both psychophysical and physiological studies have investigated how pattern vision mechanisms adapt to recent sensory history by measuring responses before and after prolonged exposure to a high-contrast visual stimulus. As was clear from the earliest experiments of this kind, adaptation elevates pattern detection thresholds in a selective manner (Blakemore & Campbell, 1969; Pantle & Sekuler, 1968). Effects of similar nature are observed in primary visual cortex (Movshon & Lennie, 1979), but not at earlier stages of visual processing, although this view has recently been challenged (Kohn, 2007; Solomon, Peirce, Dhruv, & Lennie, 2004). In general, adaptation studies strongly suggest that the neural mechanisms that mediate near-threshold pattern vision reside in V1. Because the cellular mechanisms underlying adaptation of V1 neurons are fairly well understood, we can implement neural adaptation effects in the NPM and evaluate the behavioral consequences.

In striate cortex, prolonged stimulation with an adequate stimulus affects neural encoding of visual information primarily through hyperpolarization of the cell membrane, which temporarily reduces neural responsiveness (Carandini & Ferster, 1997; Carandini, Movshon, & Ferster, 1998; Sanchez-Vives, Nowak, & McCormick, 2000). A cell will fire a spike when fluctuations in its membrane potential that result from summated synaptic input cross a precisely determined threshold value. Pattern adaptation induces a downward shift of the membrane potential; that is, it produces a tonic hyperpolarization. When adapted, fewer fluctuations will cross the spike threshold and neural responsiveness is reduced. In terms of functional models of cortical computation, hyperpolarization can be thought of as a reduction in the gain of the linear filtering stage. As can be seen in Figure 12, a reduction in linear filter gain is sufficient to capture the most prominent functional characteristics of pattern adaptation at the neural level; that is, the resulting shift of the contrast response function of single sensory units is well in line with physiological observations.

Adaptation to a stimulus that fails to drive a V1 cell usually has little effect on its functional response properties (Kohn, 2007; Movshon & Lennie, 1979). In other words, the strength of neuronal adaptation depends on the effectiveness of the adapter to drive the neuron. To capture this stimulus-selective nature of adaptation effects in the NPM, we assume that the change in filter gain is entirely determined by the responsiveness of a cell during the initial long adaptation phase of adaptation experiments. We fitted the relation between response history (i.e., accumulated spikes during the adaptation phase) and postadaptation filter gain using a decay function characterized by an inverse cumulative log-Gaussian function (see Figure 12B). Two free parameters (γ and δ) characterize the mean and variance of the cumulative log-Gaussian.

The progressive decrease in neural response strength caused by prolonged stimulation with a high-contrast stimulus typically develops over many seconds until it eventually saturates (e.g., Sanchez-Vives et al., 2000). There is thus an upper limit to the amount of adaptation that can be achieved. To capture this notion in our model, we define response history as the number of fired
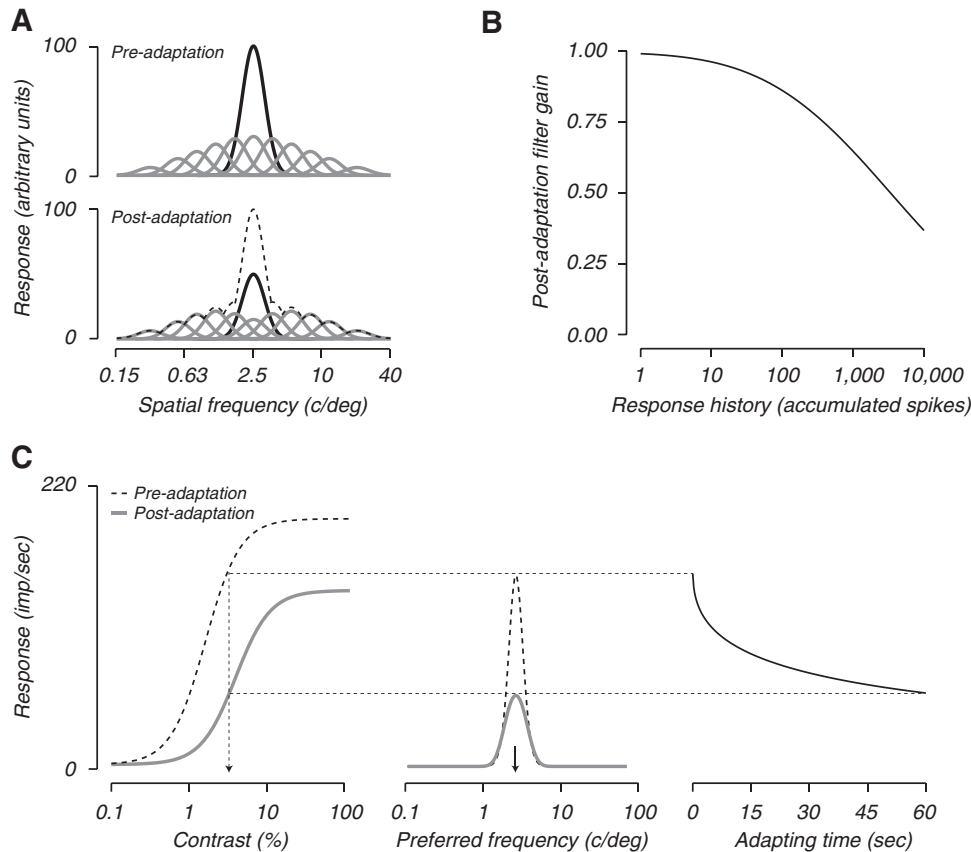
Figure 12. A simple model of cortical adaptation. (A) In the neural population model (NPM), the mechanism underlying adaptation effects is a response-history-dependent reduction in the gain of the sensory units' linear filter stage. The more spikes fired during the adaptation phase, the stronger the reduction in filter gain. This is illustrated in detail for a single sensory unit. After adaptation to the preferred stimulus, the gain of both the excitatory and inhibitory filters is (selectively) reduced. (B) The fitted relation between postadaptation filter gain and recent response history (i.e., the number of spikes fired during the adaptation phase). (C) The left panel shows the contrast response function for a sensory unit in the NPM before and after prolonged stimulation with the preferred stimulus (a high-contrast sinusoidal grating with a spatial frequency of 2.5 cycles/degree [c/deg]). After adaptation, the contrast response function is shifted to higher contrasts and lower response rates. The middle panel illustrates the full population response to a 2.5 c/deg grating at a contrast of 3% (the arrow indicates the spatial frequency of the adapter). Prolonged presentation of a stimulus reduces its effectiveness to drive the population. The right panel shows the temporal dynamics of adaptation for the sensory unit shown in the left panel.

spikes during a certain temporal interval. Free parameter ε, expressed in seconds, controls the size of the spike integration window. The temporal dynamics of adaptation in the NPM are illustrated for an example sensory unit in Figure 12C.

In sum, adaptation selectively reduces filter gain in the NPM. The strength, spread, and temporal dynamics of this effect are controlled by three free parameters (see Table 2). This implementation of adaptation effects is simple and ignores many details of the biology, but it captures the main physiological findings.

Adaptation Experiments 2: Effects of Adapting Time, Contrast, and Spatial Frequency

At the behavioral level, the main effect of pattern adaptation is a selective reduction of stimulus detectability: Upon adaptation, an observer's sensitivity for stimuli that are similar to the adapting

stimulus is diminished (Blakemore & Campbell, 1969). Not surprisingly, this general result is in line with the NPM. Adaptation selectively reduces filter gain in the model, and it follows that stimulus detectability will be compromised. But the NPM can account for various, more subtle aspects of adaptation effects as well.

The number of fired spikes during the adaptation phase determines postadaptation filter gain in the model. Consequently, the predicted reduction in stimulus visibility will depend on adapting time and contrast (more time and higher contrasts both yield more spikes). As can be seen in the upper left panel of Figure 13, the NPM prediction is in reasonable agreement with the behavioral measurements of Blakemore and Campbell (1969; RMSE = 9.6% of the relative threshold contrast). Detection thresholds initially rise fast with adapting time but saturate after about a minute. The saturation level depends on the adapting contrast. This latter effect is shown in more
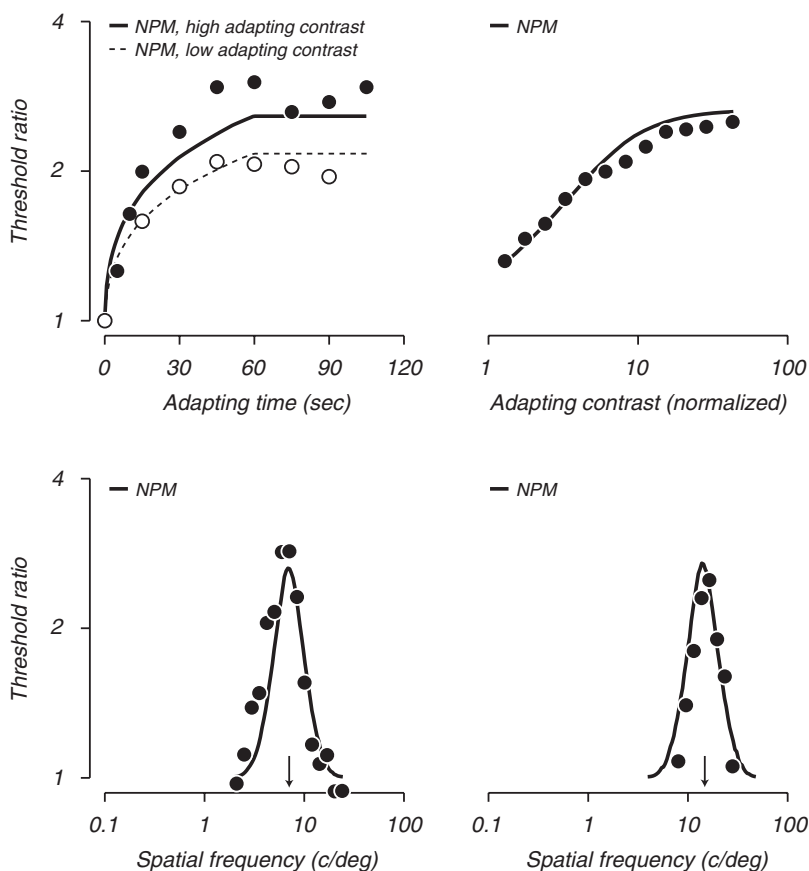
Figure 13. Threshold visibility depends on adapting time, contrast, and spatial frequency. The upper left panel illustrates how detectability of a sine wave grating selectively depends on both adapting contrast and time. White symbols refer to a low adapting contrast (0.75 log units above detection threshold), black symbols to a high adapting contrast (1.5 log units above detection threshold); the lines show the model prediction. The behavioral data are replotted from Blakemore and Campbell (1969). Both the adapting and test grating had a spatial frequency of 15 cycles/degree (c/deg). The upper right panel shows in more detail how stimulus detectability decreases with adapting contrast (normalized relative to the detection threshold) for observer DJT (black symbols) and the neural population model (NPM; thick black line). The data are replotted from Tolhurst (1972). The adapting phase in this experiment lasted for 3 min. Both the adapting and test grating had a spatial frequency of 9 c/deg. The lower panels illustrate how adaptation to a high-contrast grating (1.5 log units above detection threshold) presented for 1 min affects stimulus detectability at various spatial frequencies. The adapting grating had a spatial frequency of 7.1 c/deg (lower left panel) and 14.2 c/deg (lower right panel), respectively. The data are replotted from Blakemore and Campbell (1969).

detail in the upper right panel of Figure 13. Over almost 2 log units of adapting contrast, the model's predictions closely follow Tolhurst's (1972) data (RMSE = 5.6% of the relative threshold contrast).

Finally, consider the effects of spatial frequency The lower panels of Figure 13 illustrate how the adaptation-induced decrease in detectability depends on stimulus frequency for human observers (Blakemore & Campbell, 1969) and the NPM. Again, the model does a reasonable job in accounting for the behavioral data; that is, it produces threshold curves that are geometrically centered at the adapting frequency and have a bandwidth of approximately 1 octave at half height (RMSE = 17.5% of the relative threshold contrast).

In sum, the temporal dynamics, contrast dependency, and spectral selectivity of behavioral pattern adaptation are well explained by a model of detection in which adaptation hinders stimulus encoding through a response-history-dependent reduction of filter gain.

Adaptation Experiments 3: Of Sine Waves and Square Waves

The threshold curves shown in the lower row of Figure 13 seemingly provide strong support for the notion that adaption effects are frequency specific. However, this is not the only possible interpretation. Adaptation effects could also be period specific. One way to disentangle the effects of the period of a pattern and its frequency is to perform adaptation experiments with stimuli of more complex spectral composition, such as square wave gratings. Two variants exist: adaptation to a sine wave followed by detection of a square wave (Tolhurst, 1972) and adaptation to a square wave followed by detection of a sine wave (Blakemore & Campbell, 1969; Nachmias, Sansbury, Vassilev, & Weber, 1973; Tolhurst, 1972). In previous sections,

we have shown that the NPM accounts well for square wave detectability and sine wave adaptation. For this reason, we opted not to include the square wave adaptation experiments in the data set to which the model was fitted. The predictions of the NPM discussed here are thus blind predictions.

We first consider detection of a square wave after adaptation to a sine wave whose frequency corresponds to the square wave's fundamental. As demonstrated by Tolhurst (1972), square wave thresholds increase with adapting contrast much like sine wave thresholds do. The upper left panel of Figure 14 illustrates the prediction of the NPM, which follows the human data quite well (RMSE = 5.0% of the relative threshold contrast). The model thus provides a reasonable account for square wave detection following sine wave adaptation.

The more interesting case, however, is provided by sine wave detection after prolonged viewing of a square wave. If adaptation effects are truly mediated by frequency-tuned, independent mechanisms, one could expect to see a threshold curve that is significantly elevated at the square wave's third and fifth harmonics. This is exactly what Blakemore and Campbell (1969) found. Their data are shown in the lower panels of Figure 14. The NPM predicts the threshold elevation at the fundamental harmonic quite well (at least on average), but underestimates the elevation at the higher harmonics. This can be seen even clearer in the upper right panel of Figure 14, which plots detectability of a 9 cycles/degree sine wave after adaptation to a 3 cycles/degree square wave for a range of adapting contrasts (data replotted from Tolhurst, 1972). Again, the threshold elevation predicted by the model is considerably smaller than the human data.

In the NPM, the mild decrease of detectability around the higher harmonics is due to the inhibition provided by the units tuned to the first harmonic. During the adaptation phase, this inhibition
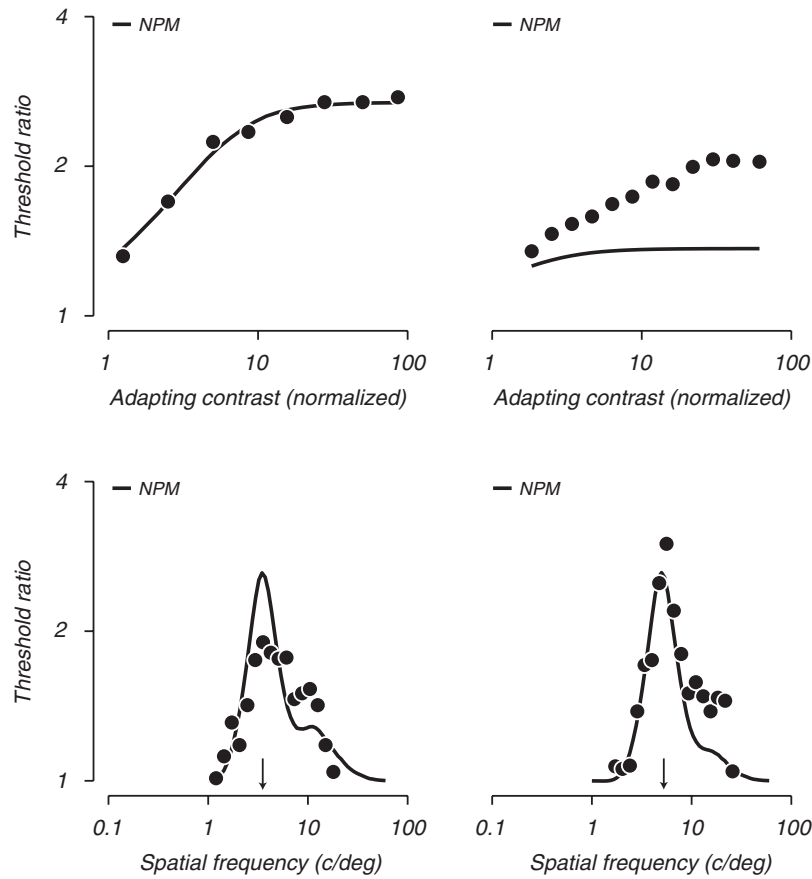

Figure 14. Adaptation experiments involving square wave gratings. The upper left panel shows how detectability of a 4.5 cycles/degree (c/deg) square wave grating decreases after adaptation to a 4.5 c/deg sine wave grating for observer DJT (black symbols) and the neural population model (NPM; thick black line). The upper right panel plots detectability of a 9 c/deg sine wave after adaptation to a 3 c/deg square wave for the same observer. The contrast of the adapting square wave is normalized with respect to the detection threshold of the second harmonic component (i.e., a 9 c/deg sinusoid). The data in the upper panels are replotted from Tolhurst (1972). The adapting phase in these experiment lasted for 3 min. The lower panels illustrate how adaptation to a high-contrast square wave grating (1.5 log units above detection threshold) presented for 1 min affects sine wave detectability at various spatial frequencies. The adapting stimulus had a spatial frequency of 3.5 c/deg (lower left panel) and 5 c/deg (lower right panel), respectively. The data are replotted from Blakemore and Campbell (1969).

limits the number of spikes fired by the units tuned to the higher harmonics. Consequently, the postadaptation filter gain remains fairly high (see Figure 11B), and stimulus detectability is not hindered very much, in contrast with the behavioral data (see Figure 14). This miss may well be a failure of the model. However, caution is warranted in interpreting this result. Several studies have tried to replicate the psychophysical result but failed to do so. Nachmias et al. (1973) and Klein and Stromeyer (1980) used a 2AFC procedure instead of the method of adjustment that was employed in the earlier studies. Both studies convincingly replicated the loss of detectability around the first harmonic but could not find a significant threshold rise around the third harmonic, much like predicted by the NPM.

In this section, we have demonstrated that the NPM economically accounts for several behavioral adaptation results involving sine wave gratings. When the adapter is a square wave grating, the model predicts only a small threshold increase around the higher harmonics. This result is consistent with some, but not all, psychophysical reports.

Uncertainty Experiments 1: Detection of Signals With Varying Prior Probability

The final series of experiments we discuss emphasize the role of decoding in visual information processing. The ubiquity of sensory noise throughout cortex poses a serious challenge for perceptual decision making. If an organism is to perform well in sensory tasks, it must tactfully ignore the huge number of sensory re-

sponses that occur at any point in time without conveying information of direct interest. Indeed, sometimes it may even be advantageous for an observer to conclude that the available sensory input is unreliable and instead largely base his or her judgment on prior knowledge of the relevant statistics for the task at hand. It is this type of perceptual computation that allows us to localize lost pacifiers in baby cribs in the dark of the night. Bayesian statistics provide a normative framework that describes how noisy measurements and prior information ought to be combined to optimally achieve a predefined goal. Recent years have seen an avalanche of research dedicated to the mechanisms that govern perceptual decision making. There is now a large body of evidence suggesting that human sensory computations often resemble Bayes's optimal (Knill & Pouget, 2004) or at least are consistent with the notion that behavioral performance approaches the optimum that can be achieved in a noisy system subject to certain constraints (Geisler, 2008). The model we propose here resonates well with this view. Sensory responses are read out by a linear decoder that aims to maximize task performance without knowing the details of the full population response distributions. Specifically, we assume that each neuron's contribution to the decision statistic is determined by a simple model of its reliability to perform the task at hand (Equation 9), without taking the interneuronal noise correlations into account. In all experiments discussed thus far, the diverse stimulus manipulations affected the noisy sensory measurements but not the prior statistics of the task. We now turn our attention to a paradigm that examines the effects
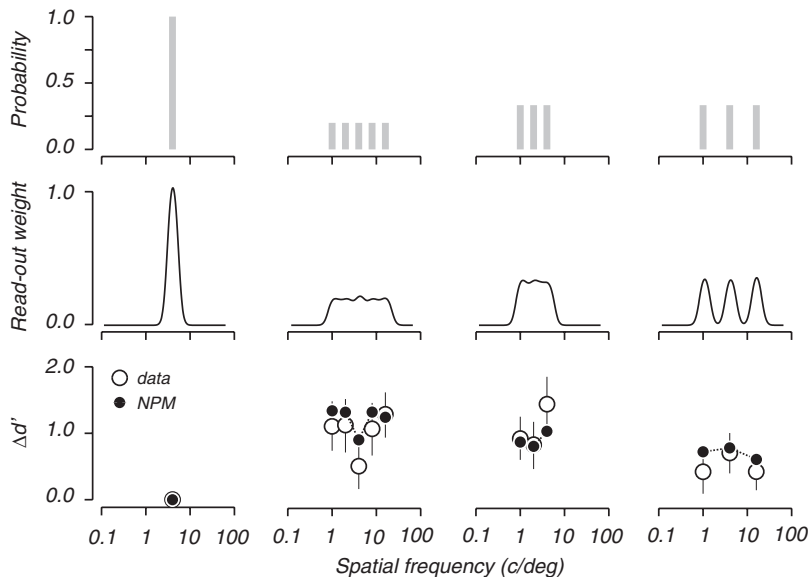

Figure 15. Detection of multiple equally probable signals. The upper row illustrates stimulus probability in a signal-known-exactly experiment (left) and three uncertainty experiments. In the neural population model (NPM), these signal probabilities give rise to different read-out strategies employed by the decoder. Read-out weights are plotted in the middle row. The predicted loss in performance, together with data obtained from human observers, is shown in the lower row (observers TS [column 2], PK [column 3], and TS [column 4]; data replotted from Davis et al., 1983). Error bars reflect the 68% confidence interval. Davis et al. (1983) plotted these data in percentage correct. To estimate the corresponding d' values, we assumed that the observers in the two-alternative forced-choice uncertainty experiment were unbiased in their decision making. To obtain confidence intervals on the d' estimates, we assumed binomially distributed choice responses. c/deg = cycles/degree.

of changing the prior but not the stimulus: uncertainty experiments.

Uncertainty experiments provide an elegant means to investigate to what degree human pattern detection reflects task-specific decoding. In these studies, observers have to detect a pattern while being either certain or uncertain about some aspect of the stimulus. For instance, on any given trial, the signal frequency is either known to the observer or can be one of multiple possible frequencies. In Bayesian terms, uncertainty experiments compare conditions with different prior distributions of the signal to be detected. In the case of spatial frequency, this manipulation affects pattern visibility: Broader priors (more uncertainty) reduce stimulus detectability (Davis, 1981; Davis & Graham, 1981; Davis, Kramer, & Graham, 1983).

In the NPM, uncertainty manipulations affect the read-out weights employed by the decoder in a deterministic fashion. For this reason, we did not include the uncertainty experiments in the data set to which the model was fitted. Instead, we compare a blind model prediction to the empirical findings of seven uncertainty experiments performed by Davis (1981), Davis and Graham (1981), and Davis et al. (1983).

It is straightforward to derive model predictions for an uncertainty experiment in the NPM. A maximum-likelihood decoder fully exploits its knowledge about task characteristics. It knows which neurons contribute information and which can be safely ignored. For this decoder, the optimal strategy when uncertain about stimulus frequency consists of linearly combining the weights optimized to the signals presented in isolation whereby the combinatorial weights reflect stimulus probability. Let w_1 and w_2 be the set of weights optimized to detect S_1 and S_2, respectively. In the intermixed condition of the uncertainty experiment, $P(S_1) = q$ and $P(S_2) = 1 - q$. The optimal weights w_{mix} are given by

$$w_{mix} = qw_1 + (1 - q)w_2. \qquad (15)$$

Our decoder thus takes the full prior probability distribution into account. On any given trial, the resulting weighting profile w_{mix} includes all potentially informative neurons as well as many noninformative neurons that contribute noise to the decision statistic. Consequently, detection performance drops relative to the signal-known-exactly condition. This is illustrated in Figure 15 for three experiments that measure detectability of multiple equally probable signals (Davis et al., 1983). Note that the predictions of the NPM vary as a function of the size and range of the stimulus ensemble: Uncertainty effects are maximal for large stimulus ensembles that span a wide range of frequencies. Although the human data lack the reliability to assess whether they differ across experiments, they are generally in good agreement with the blind model predictions.

In a different series of experiments, Davis (1981) and Davis and Graham (1981) sought to determine how detection performance is affected by mixing multiple signals of varying probability. For instance, one of their observers had to detect two far-apart frequencies that occur with a probability of 47.5% and three in-between frequencies that occur on only 1.67% of the trials. In the NPM, such manipulations naturally give rise to a tuned uncertainty effect: The loss in performance is most dramatic for signals of low probability that are far removed from highly probable signals. As can be seen in Figure 16, this prediction is again well in line with the behavioral findings.

Finally, we note that the effectiveness of stimulus uncertainty manipulations depends on the stimulus dimension considered. Manipulations involving uncertainty about signal contrast, for instance, hardly affect signal detectability for human observers (Davis et al., 1983). Interestingly, this is to be expected from the NPM as well. Contrast gain control renders neural tuning functions largely contrast invariant (Finn, Priebe, & Ferster, 2007; Geisler & Albrecht, 1995). Consequently, the weighting
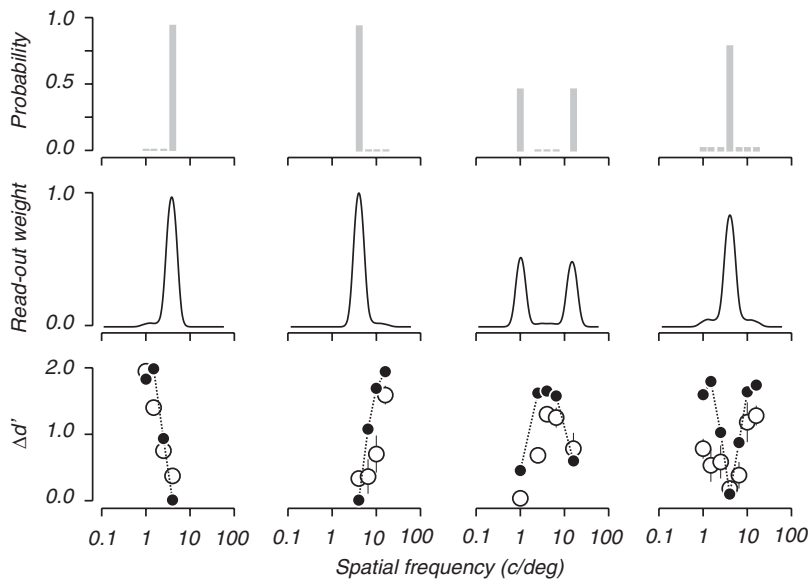

Figure 16. Detection of multiple signals of varying probability. Plot conventions are the same as in Figure 15. The data are from observers TS, ED, ED (replotted from Davis, 1981), and TS (replotted from Davis & Graham, 1981), respectively. c/deg = cycles/degree.

profile employed by the maximum-likelihood decoder does not depend much on stimulus contrast either. There is no appreciable uncertainty effect for stimulus contrast in the NPM.

Uncertainty Experiments 2: Uncertainty and the Slope of the Psychometric Function

Uncertainty effects have often been considered in the context of intrinsic uncertainty models. According to these models, the fundamental limit in detection is provided by inefficient decision making; that is, observers base their decision on a single sensory response but fail to distinguish informative from noninformative sensory responses (Lu & Dosher, 2008; Pelli, 1985). The higher the stimulus contrast, the more likely that the relevant channel provides the maximal response. Not surprisingly, this strategy leads to an increase in thresholds with growing uncertainty. In addition, uncertainty affects the slope of the psychometric function (Pelli, 1985): The higher the uncertainty, the steeper the slope of the psychometric function on semilogarithmic coordinates (see Figure 17).

This is not the case in the NPM. Consider the weighting profiles corresponding to low, medium, and high uncertainty in Figure 17. On every single trial, the relevant sensory units will make some contribution to the decision statistic. When uncertainty grows, the relative weight of that contribution decreases. In addition, the weight assigned to uninformative units increases. Increasing uncertainty thus effectively leads to scaling down the stimulus contrast and increasing the background noise. However, because the weighting profiles are contrast independent, the relative reduction in signal-to-noise ratio is contrast independent too. In other words, in the NPM, the slope of the psychometric function on logarithmic contrast coordinates does not depend on uncertainty. This is illustrated in Figure 17. To test this prediction, one could measure full psychometric functions under conditions of low and high un-

certainty (the larger the threshold increase, the better). To the best of our knowledge, such experimental data are not yet available.

In this section, we have argued that human observers rely on the stimulus probability distributions that define a detection task in a manner that is well described by a reduced Bayesian decoder directly operating on the raw sensory responses at the level of V1. Davis et al. (1983) already demonstrated unambiguously that detection strategies vary with stimulus probability. The NPM now offers a principled means to quantify this insight in a parsimonious way—results from all seven uncertainty experiments were *predicted* by our model, not fitted. In contrast to earlier models, the NPM predicts that the slope of the psychometric function does not depend on stimulus certainty.

Discussion

In the introduction of this article, we motivated the need for a new computational theory of visual pattern detection and identified the following set of desired characteristics, all related to the general theme of unification: first, an explicit connection with the neurophysiological underpinnings of pattern vision; second, a direct link with the mechanisms involved in pattern masking; third, a perceptual decision-making strategy that generalizes beyond detection; and fourth, applicability to a broad range of behavioral results. We introduced and evaluated the neural population model, designed to meet these goals. This theory builds on several earlier models and findings in the vision literature (e.g., Blakemore & Campbell, 1969; Campbell & Robson, 1968; Davis et al., 1983; Dosher & Lu, 1998; Foley, 1994; Geisler & Albrecht, 1997; Graham & Nachmias, 1971; Itti et al., 2000; Legge & Foley, 1980; Parker & Newsome, 1998; Pelli, 1985; Petrov et al., 2005). It is a mechanistic model in which visual information is encoded by a neurophysiologically inspired front-end resembling V1 and decoded by a deci-
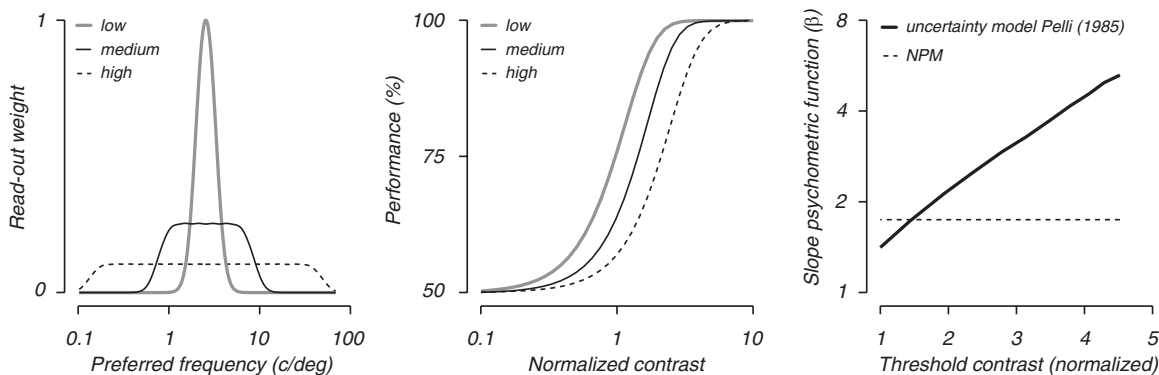

Figure 17. Uncertainty and the slope of the psychometric function. The left panel illustrates the weighting profiles for three extrinsic uncertainty conditions in the neural population model (NPM): low, medium, and high uncertainty. Each profile is geometrically centered at 2.5 cycles/degree (c/deg). The middle panel shows the corresponding psychometric functions for a sinusoidal grating of 2.5 c/deg. Increasing uncertainty reduces stimulus detectability but does not alter the slope of the psychometric function on semilogarithmic coordinates in the NPM. The beta parameter of the psychometric function is thus constant for the NPM. This result is in contradiction with a well-known model for uncertainty effects, discussed in detail by Pelli (1985). In this model, the slope of the psychometric function covaries with its threshold, such that more uncertainty produces steeper psychometric functions (right panel).

sion maker that takes into account the reliability of individual neurons in performing the task at hand. Behavioral performance in the model is limited by the presence of correlated sensory noise. Using a single set of parameters, the model successfully explains a large body of findings in the detection literature, thereby offering a new interpretation for several classic results in summation, adaptation, and uncertainty experiments.

As a case in point, we consider summation in more detail. Summation studies have investigated to what degree threshold vision obeys homogeneity, additivity, and superposition. In the NPM, these experiments drive a complex combination of mechanisms that involve selective filtering, response squaring, broadly tuned response suppression, and task-specific response pooling. Acting in concert, these mechanisms give a remarkably successful account of summation results. For instance, Graham and Nachmias (1971) made the well-known observation that adding a low-contrast grating to a threshold sinusoid hardly affects its visibility when both gratings reasonably differ in spatial frequency. In the NPM, this manipulation reduces sensory response strengths through gain control (see Figure 8). However, this response reduction is compensated by the maximum-likelihood decoder that at the same time exploits the increase in the absolute number of responsive units, producing the net null result (see Figure 9).

The main benefit of this alternative account of summation is that it naturally extends to findings in pattern masking. For instance, adding a second low-contrast grating to the signal and the background is known to improve stimulus detectability profoundly, provided that both gratings are reasonably close in spatial frequency and orientation (Legge & Foley, 1980; Nachmias & Sansbury, 1974). This nonlinear effect is largely driven by response expansion and therefore occurs in the NPM as well (see Figure 18). On the other hand, adding a high-contrast grating to the signal and the background greatly reduces stimulus detectability, even when signal and mask differ considerably (Foley, 1994; Legge & Foley, 1980). This masking effect is testament to the role of a broadly tuned gain control mechanism and thus also consistent with the NPM (see Figure 18). In sum, the NPM explains threshold summation using the very mechanisms identified in suprathreshold studies and thus explicitly connects both fields.

The success of the various multichannel models that make use of linear–nonlinear cascades and population decoding to account for important aspects of detection, pattern masking (e.g., Itti et al., 2000), visual attention (e.g., Lee et al., 1999), and perceptual learning (e.g., Petrov et al., 2005) raises the question whether it is possible to formulate a single model that achieves all of this at once. Given that the proposed models differ in many of their
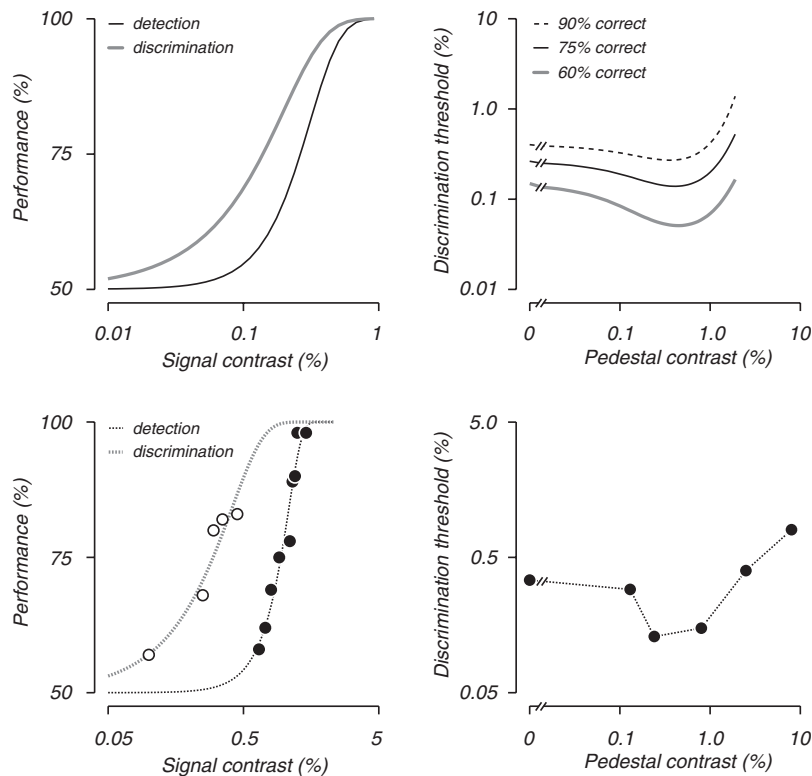

Figure 18. Contrast discrimination. The upper left panel illustrates the model's psychometric function for detection and discrimination (whereby the pedestal contrast is equal to the detection threshold). At this pedestal contrast, discrimination is better than detection, but the magnitude of the performance difference depends on the performance level considered. The upper right panel shows how discrimination thresholds computed at three performance levels evolve with pedestal contrasts in the neural population model. The lower panels illustrate some of the measurements performed by Nachmias and Sansbury (1974) for 9 cycles/degree gratings (left) and 3 cycles/degree gratings (right), respectively. Discrimination thresholds were computed at a performance level of 79.4% correct.

details (e.g., the mechanisms underlying decision making), this is not self-evident. Furthermore, taking the large number of empirical phenomena that need consideration into account, it seems likely that such project will turn out to be a major undertaking. Nevertheless, we believe that unifying different subfields in early vision is a goal well worth pursuing.

We now draw attention to the decision-making strategy implemented in our model. The encoding stage in the NPM solely makes use of canonical neural computations: linear filtering, response expansion, and normalization (Carandini & Heeger, 2012). At first sight, the decoder requires more complex arithmetic, as it needs to estimate response difference reliabilities (Equation 9). We wish to emphasize that this impression is misleading. The decoder assumes neural noise to be normally distributed and statistically independent. Both assumptions are simplifications and wrong in their details. However, they enable implementing a maximum-likelihood inference strategy using nothing more than simple weighted summation of the input activities. To see this, bear in mind that cortical response variances are proportional to the mean response and that rescaling all weights with the same factor does not affect network behavior. Therefore,

$$w_i = \frac{E[r(s) - r_0]}{\text{VAR}(r(s) - r_0)} \qquad (9)$$

can be rewritten as

$$w_i = \frac{E[r(s) - r_0]}{E[r(s) + r_0]}. \qquad (16)$$

The fact that the model is invariant to rescaling all read-out weights with a common scalar has the additional benefit that it allows to circumvent the problem that neurons representing summed activities may saturate. This can be done by incorporating a shared normalization signal. Thus, the decoder employed in our study can be realized in the same neural machinery used to encode visual information in the model, that is, a linear–nonlinear cascade in which weighted summation is followed by divisive inhibition.

Relation With the Linear Channels Model

The model for pattern detection that we advocated in this article builds directly on the foundations provided by the early theoretical framework. Although the NPM differs considerably in its characterization of the putative visual mechanisms that mediate visual encoding and decoding, it is in full agreement with the two most innovative claims of the linear channels model. First, visual encoding in our model begins with a local Fourier analysis, performed by the linear operator that defines the sensory units' stimulus selectivity. This insight remains one of the most powerful contributions ever made to the vision literature. Second, the encoded visual information is not further processed by higher visual areas but directly linked to perceptual judgments. In our model, the quality of these judgments is ultimately limited by correlated neural noise. As we pointed out, this limitation enables the NPM to accurately describe the relation between neural and behavioral sensitivity (see Figure 6), a rare feat for models of psychophysical tasks. An additional consequence is that the NPM predicts significant trial-to-trial correlations between neuronal activity in V1 and choice behavior in a detection task (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Nienborg & Cumming, 2010; Parker & Newsome, 1998).[6] A recent study involving macaque monkeys performing pattern detection reported exactly this: significant choice probabilities in V1 (Palmer et al., 2007). Thus, when a monkey is engaged in a detection task and repeatedly is confronted with the same difficult stimulus, fluctuations in single cell responses in V1 carry information about the behavioral outcome of this choice process. As yet, this finding is the most direct physiological evidence that pattern detection is mediated by the sensory computations performed in primary visual cortex. Its discovery and consistency with the PTM lends strong support to the bold ambition that was first expressed in the linear channels model, that we can directly connect visual behavior to its physiological substrate using quantitative information processing theories as the bridge.

---

[6] This correlation is a statistic that is often used in awake animal neurophysiology to study the neural correlates of perception. It is termed "choice probability" and expresses to which extent the trial-to-trial variability in the neuronal activity of a sensory brain area is reflected in an animal's behavioral choices (Britten et al., 1996).

References

Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In M. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 3–20). Cambridge, MA: MIT Press.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience, 7,* 358–366. doi:10.1038/nrn1888

Baker, D. H., & Meese, T. S. (2012). Zero-dimensional noise: The best mask you never saw. *Journal of Vision, 12,* 20. doi:10.1167/12.10.20

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception, 1,* 371–394. doi:10.1068/p010371

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron, 74,* 30–39. doi:10.1016/j.neuron.2012.03.016

Bex, P. J., Solomon, S. G., & Dakin, S. C. (2009). Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure. *Journal of Vision, 9,* 1. doi:10.1167/9.10.1

Bird, C. M., Henning, G. B., & Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 19,* 1267–1273. doi:10.1364/JOSAA.19.001267

Blackwell, K. T. (1998). The effect of white and filtered noise on contrast detection thresholds. *Vision Research, 38,* 267–280. doi:10.1016/S0042-6989(97)00130-2

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203,* 237–260.

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience, 13,* 87–100. doi:10.1017/S095252380000715X

Campbell, F. W., Carpenter, R. H. S., & Levinson, J. Z. (1969). Visibility of aperiodic patterns compared with that of sinusoidal gratings. *Journal of Physiology, 204,* 283–298.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551–566.

Carandini, M., & Ferster, D. (1997). A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. *Science, 276,* 949–952. doi: 10.1126/science.276.5314.949

Carandini, M., & Heeger, D. J. (2012). Normalization as canonical neural computation. *Nature Reviews Neuroscience, 13,* 51–62. doi:10.1038/nrn3136

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17,* 8621–8644.

Carandini, M., Movshon, J. A., & Ferster, D. (1998). Pattern adaptation and cross-orientation interactions in the primary visual cortex. *Neuropharmacology, 37,* 501–511. doi:10.1016/S0028-3908(98)00069-0

Chen, Y., Geisler, W. S., & Seidemann, E. (2006). Optimal decoding of correlated neural population responses in the primate visual cortex. *Nature Neuroscience, 9,* 1412–1420. doi:10.1038/nn1792

Chirimuuta, M., & Tolhurst, D. J. (2005). Does a Bayesian model of V1 contrast coding offer a neurophysiological account of human contrast discrimination? *Vision Research, 45,* 2943–2959. doi:10.1016/j.visres.2005.06.022

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., . . . Shenoy, K. V. (2010). Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nature Neuroscience, 13,* 369–378. doi:10.1038/nn.2501

Cohen, M. R., & Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature Neuroscience, 14,* 811–819. doi:10.1038/nn.2842

Cohen, M. R., & Newsome, W. T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *Journal of Neuroscience, 29,* 6635–6648. doi:10.1523/JNEUROSCI.5179-08.2009

Dao, D. Y., Lu, Z.-L., & Dosher, B. A. (2006). Adaptation to sine-wave gratings selectively reduces the contrast gain of the adapted stimuli. *Journal of Vision, 6,* 6. doi:10.1167/6.7.6

Davis, E. T. (1981). Allocation of attention: Uncertainty effects when monitoring one or two visual gratings of noncontiguous spatial frequencies. *Perception & Psychophysics, 29,* 618–622. doi:10.3758/BF03207381

Davis, E. T., & Graham, N. (1981). Spatial frequency uncertainty effects in the detection of sinusoidal gratings. *Vision Research, 21,* 705–712. doi:10.1016/0042-6989(81)90079-1

Davis, E. T., Kramer, P., & Graham, N. (1983). Uncertainty about spatial frequency, spatial position or contrast of visual patterns. *Perception & Psychophysics, 33,* 20–28. doi:10.3758/BF03205862

Dosher, B. A., & Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America, 95,* 13988–13993. doi:10.1073/pnas.95.23.13988

Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., & Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science, 327,* 584–587. doi:10.1126/science.1179867

Eckstein, M. P., Ahumada, A. J., & Watson, A. B. (1997). Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 14,* 2406–2419. doi:10.1364/JOSAA.14.002406

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433. doi:10.1038/415429a

Fetsch, C. R., Pouget, A., DeAngelis, A. G. C., & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience, 15,* 146–154. doi:10.1038/nn.2983

Finn, I. M., Priebe, N. J., & Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron, 54,* 137–152. doi:10.1016/j.neuron.2007.02.029

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 11,* 1710–1719. doi:10.1364/JOSAA.11.001710

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192. doi:10.1146/annurev.psych.58.110405.085632

Geisler, W. S., & Albrecht, D. G. (1995). Bayesian analysis of identification performance in monkey visual cortex: Nonlinear mechanisms and stimulus certainty. *Vision Research, 35,* 2723–2730. doi:10.1016/0042-6989(95)00029-Y

Geisler, W. S., & Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination and identification. *Visual Neuroscience, 14,* 897–919. doi:10.1017/S0952523800011627

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience, 14,* 926–932. doi:10.1038/nn.2831

Gorea, A., & Tyler, C. W. (1986). New look at Bloch's law for contrast. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 3,* 52–61. doi:10.1364/JOSAA.3.000052

Goris, R. L. T., Wagemans, J., & Wichmann, F. A. (2008). Modelling contrast discrimination data suggest both the pedestal effect and stochastic resonance to be caused by the same mechanism. *Journal of Vision, 8,* 17. doi:10.1167/8.15.17

Goris, R. L. T., Wichmann, F. A., & Henning, G. B. (2009). A neurophysiologically plausible population code model for human contrast discrimination. *Journal of Vision, 9,* 15. doi:10.1167/9.7.15

Goris, R. L. T., Zaenen, P., & Wagemans, J. (2008). Some observations on contrast detection in noise. *Journal of Vision, 8,* 4. doi:10.1167/8.9.4

Graf, A. B. A., Kohn, A., Jazayeri, M., & Movshon, J. A. (2011). Decoding the activity of neural populations in macaque primary visual cortex. *Nature Neuroscience, 14,* 239–245. doi:10.1038/nn.2733

Graham, N. V. S. (1989). *Visual pattern analyzers.* New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195051544.001.0001

Graham, N. (1992). Breaking the visual stimulus into parts. *Current Directions in Psychological Science, 1,* 55–61. doi:10.1111/1467-8721.ep11509742

Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision Research, 51,* 1397–1430. doi:10.1016/j.visres.2011.02.007

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single channel and multichannel models. *Vision Research, 11,* 251–259. doi:10.1016/0042-6989(71)90189-1

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York, NY: Wiley.

Heeger, D. J. (1992a). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9,* 427–443. doi:10.1017/S095252380001124X

Heeger, D. J. (1992b). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9,* 181–197. doi:10.1017/S0952523800009640

Henning, G. B. (1988). Spatial-frequency tuning as a function of temporal frequency and stimulus motion. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 5,* 1362–1373. doi:10.1364/JOSAA.5.001362

Henning, G. B., & Wichmann, F. A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7,* 3. doi:10.1167/7.1.3

Hubel, D. H. (1959). Single unit activity in striate cortex of unrestrained cats. *Journal of Physiology, 147,* 226–238.

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

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

Itti, L., Koch, C., & Braun, J. (2000). Revisiting spatial vision: Towards a unifying model. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 17,* 1899–1917. doi:10.1364/JOSAA.17.001899

Klein, S. A., & Levi, D. M. (2009). Stochastic model for detection of signals in noise. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 26,* B110–B126. doi:10.1364/JOSAA.26.00B110

Klein, S., & Stromeyer, C. F., III. (1980). On inhibition between spatial frequency channels: Adaptation to complex gratings. *Vision Research, 20,* 459–466. doi:10.1016/0042-6989(80)90037-1

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation for perception and action. *Trends in Neurosciences, 27,* 712–719. doi:10.1016/j.tins.2004.10.007

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97,* 3155–3164. doi:10.1152/jn.00086.2007

Kohn, A., & Smith, M. A. (2005). Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *Journal of Neuroscience, 25,* 3661–3673. doi:10.1523/JNEUROSCI.5106-04.2005

Law, C.-T., & Gold, J. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience, 11,* 505–513. doi:10.1038/nn2070

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience, 2,* 375–381. doi:10.1038/7286

Legge, G. E. (1978). Sustained and transient mechanisms in human vision: Temporal and spatial properties. *Vision Research, 18,* 69–81. doi:10.1016/0042-6989(78)90079-2

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471. doi:10.1364/JOSA.70.001458

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review, 115,* 44–82. doi:10.1037/0033-295X.115.1.44

Luntinen, O., Rovamo, J., & Naesaenen, R. (1995). Modelling the increase of contrast sensitivity with grating area and exposure time. *Vision Research, 35,* 2339–2346. doi:10.1016/0042-6989(94)00309-A

Macke, J. H., & Wichmann, F. A. (2010). Estimating predictive stimulus features from psychophysical data: The decision image technique applied to human faces. *Journal of Vision, 10,* 22. doi:10.1167/10.5.22

Mamassian, P., & Landy, M. S. (2001). Interaction of visual prior constraints. *Vision Research, 41,* 2653–2668. doi:10.1016/S0042-6989(01)00147-X

Morgenstern, Y., & Elder, J. H. (2012). Local visual energy mechanisms revealed by detection of global patterns. *Journal of Neuroscience, 32,* 3679–3696. doi:10.1523/JNEUROSCI.3881-11.2012

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117–151). Rome, Italy: Vatican Press.

Movshon, J. A., & Lennie, P. (1979). Pattern selective adaptation in striate cortical neurones. *Nature, 278,* 850–852. doi:10.1038/278850a0

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978a). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology, 283,* 79–99.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978b). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology, 283,* 53–77.

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039–1042. doi:10.1016/0042-6989(74)90175-8

Nachmias, J., Sansbury, R. V., Vassilev, A., & Weber, A. (1973). Adaptation to square-wave gratings: In search of the elusive third harmonic. *Vision Research, 13,* 1335–1342. doi:10.1016/0042-6989(73)90209-5

Nagaraja, N. S. (1964). Effects of luminance noise on contrast thresholds. *Journal of the Optical Society of America, 54,* 950–955. doi:10.1364/JOSA.54.000950

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature, 434,* 387–391. doi:10.1038/nature03390

Nienborg, H., & Cumming, B. (2010). Correlations between the activity of sensory neurons and behavior: How much do they tell us about a neuron's causality? *Current Opinion in Neurobiology, 20,* 376–381. doi:10.1016/j.conb.2010.05.002

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell field properties by learning a sparse code for natural images. *Nature, 381,* 607–609. doi:10.1038/381607a0

Palmer, C., Chen, Y., & Seidemann, E. (2007). Linking neuronal and behavioral performance in a reaction-time visual detection task. *Journal of Neuroscience, 27,* 8122–8137. doi:10.1523/JNEUROSCI.1940-07.2007

Pantle, A., & Sekuler, R. W. (1968). Size-detecting mechanisms in human vision. *Science, 162,* 1146–1148. doi:10.1126/science.162.3858.1146-a

Parker, A. J., Krug, K., & Cumming, B. G. (2002). Neuronal activity and its link with the perception of multi-stable figures. *Philosophical Transactions of the Royal Society B: Biological Sciences, 357,* 1053–1062. doi:10.1098/rstb.2002.1112

Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience, 21,* 227–277. doi:10.1146/annurev.neuro.21.1.227

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 2,* 1508–1532. doi:10.1364/JOSAA.2.001508

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental channel reweighting. *Psychological Review, 112,* 715–743. doi:10.1037/0033-295X.112.4.715

Pouget, A., Dayan, P., & Zemel, R. S. (2003). Inference and computation with population codes. *Annual Review of Neuroscience, 26,* 381–410. doi:10.1146/annurev.neuro.26.041002.131112

Purushothaman, G., & Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience, 8,* 99–106. doi:10.1038/nn1373

Putzeys, T., Bethge, M., Wichmann, F. A., Wagemans, J., & Goris, R. L. T. (2012). A new perceptual bias reveals suboptimal population decoding of sensory responses. *PLoS Computational Biology, 8*(4), e1002453. doi:10.1371/journal.pcbi.1002453

Rehn, M., & Sommer, F. T. (2007). A network that uses few active neurones to code visual input predicts the diverse shapes of cortical receptive fields. *Journal of Computational Neuroscience, 22,* 135–146. doi:10.1007/s10827-006-0003-9

Sachs, M. B., Nachmias, J., & Robson, J. G. (1971). Spatial frequency channels in human vision. *Journal of the Optical Society of America, 61,* 1176–1186. doi:10.1364/JOSA.61.001176

Sanborn, A. N., & Dayan, P. (2011). Optimal decisions for contrast discrimination. *Journal of Vision, 11,* 9. doi:10.1167/11.14.9

Sanchez-Vives, M. V., Nowak, L. G., & McCormick, D. A. (2000). Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. *Journal of Neuroscience, 20,* 4267–4285.

Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience, 16,* 1486–1510.

Solomon, S. G., Peirce, J. W., Dhruv, N. T., & Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron, 42,* 155–162. doi:10.1016/S0896-6273(04)00178-3

Stromeyer, C. F., III, & Klein, S. (1975). Evidence against narrow-band spatial frequency channels in human vision: The detectability of frequency modulated gratings. *Vision Research, 15,* 899–910. doi:10.1016/0042-6989(75)90229-1

Tanner, W. P., Jr., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61,* 401–409. doi:10.1037/h0058700

Tolhurst, D. J. (1972). Adaptation to square-wave gratings: Inhibition between spatial frequency channels in the human visual system. *Journal of Physiology, 226,* 231–248.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23,* 775–785. doi:10.1016/0042-6989(83)90200-6

Wandell, B. A. (1995). *Foundations of vision*. Sunderland, MA: Sinauer.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598–604. doi:10.1038/nn0602-858

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: Contrast discrimination* (Unpublished doctoral dissertation). University of Oxford, Oxford, England.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63,* 1293–1313. doi:10.3758/BF03194544

Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63,* 1314–1329. doi:10.3758/BF03194545

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature, 370,* 140–143. doi:10.1038/370140a0

Appendix

Experimental Methods

Here we describe the experimental methods of the newly conducted summation experiment in detail. Stimuli were presented with 8-bit luminance precision on a calibrated monochrome Siemens SMM 21106 LS monitor with a white phosphor (P-45) and spatial resolution of 996 × 777 pixels at 130 Hz. Targets were weighted sums of a 3 and 9 cycles degree^{-1} sine wave grating with randomly selected phases, windowed by a Gaussian envelope (σ = 1.27° of visual angle). All stimuli were orientated horizontally. Background luminance was 50 cd m^{-2}.

On each trial observers first fixated a cross at the center of the display for 250 ms. The cross disappeared 500 ms before the onset of two auditory cued 50-ms intervals, separated by 500 ms. One interval contained the target, presented at one of seven contrasts, and the other contained a blank. Observers indicated which interval they thought contained the target. Auditory feedback was provided after each trial. Within a block of 50 trials, the target was held constant while its contrast was randomized. Observers completed at least 4,900 trials (seven targets × seven contrasts × 100 trials). The data for each target were fit with a Weibull function. The 75% correct threshold was determined with maximum-likelihood methods (Wichmann & Hill, 2001a), its confidence interval with Monte Carlo parametric bootstrap procedures (Wichmann & Hill, 2001b).