

Do We Know What the Early Visual System Does?

Matteo Carandini,¹ Jonathan B. Demb,² Valerio Mante,¹ David J. Tolhurst,³ Yang Dan,⁴ Bruno A. Olshausen,⁶ Jack L. Gallant,^{5,6} and Nicole C. Rust⁷

¹Smith-Kettlewell Eye Research Institute, San Francisco, California 94115, ²Departments of Ophthalmology and Visual Sciences, and Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48105, ³Department of Physiology, University of Cambridge, Cambridge CB2 1TN, United Kingdom, Departments of ⁴Molecular and Cellular Biology and ⁵Psychology and ⁶Helen Wills Neuroscience Institute and School of Optometry, University of California, Berkeley, Berkeley, California 94720, and ⁷Center for Neural Science, New York University, New York, New York 10003

We can claim that we know what the visual system does once we can predict neural responses to arbitrary stimuli, including those seen in nature. In the early visual system, models based on one or more linear receptive fields hold promise to achieve this goal as long as the models include nonlinear mechanisms that control responsiveness, based on stimulus context and history, and take into account the nonlinearity of spike generation. These linear and nonlinear mechanisms might be the only essential determinants of the response, or alternatively, there may be additional fundamental determinants yet to be identified. Research is progressing with the goals of defining a single “standard model” for each stage of the visual pathway and testing the predictive power of these models on the responses to movies of natural scenes. These predictive models represent, at a given stage of the visual pathway, a compact description of visual computation. They would be an invaluable guide for understanding the underlying biophysical and anatomical mechanisms and relating neural responses to visual perception.

Key words: contrast; lateral geniculate nucleus; luminance; primary visual cortex; receptive field; retina; visual system; natural images

The ultimate test of our knowledge of the visual system is prediction: we can say that we know what the visual system does when we can predict its response to arbitrary stimuli. How far are we from this end result? Do we have a “standard model” that can predict the responses of at least some early part of the visual system, such as the retina, the lateral geniculate nucleus (LGN), or primary visual cortex (V1)? Does such a model predict responses to stimuli encountered in the real world?

A standard model existed in the early decades of visual neuroscience, until the 1990s: it was given by the linear receptive field. The linear receptive field specifies a set of weights to apply to images to yield a predicted response. A weighted sum is a linear operation, so it is simple and intuitive. Moreover, linearity made the receptive field mathematically tractable, allowing the fruitful marriage of visual neuroscience with image processing (Robson, 1975) and with linear systems analysis (De Valois and De Valois, 1988). It also provided a promising parallel with research in visual perception (Graham, 1989). Because it served as a standard model, the receptive field could be used to decide which findings were surprising and which were not: if a phenomenon was not predictable from the linear receptive field, it was particularly worthy of publication.

Research aimed at testing the linear receptive field led to the discovery of important nonlinear phenomena, which cannot be explained by a linear receptive field alone. These phenomena

have been discovered at all stages of the early visual system, including the retina (for review, see Shapley and Enroth-Cugell, 1984; Demb, 2002), the LGN (for review, see Carandini, 2004), and area V1 (for review, see Carandini et al., 1999; Fitzpatrick, 2000; Albright and Stoner, 2002). They have forced a revision of the models based on the linear receptive field. In some cases, the revised models have achieved near standard model status, for example, the model of Shapley and Victor for contrast gain control in retinal ganglion cells (Shapley and Victor, 1978; Victor, 1987) and Heeger’s normalization model of V1 responses (Heeger, 1992a). By and large, however, the discovery of failures of the linear receptive field has deprived the field of a simple standard model for each visual stage.

This review aims to help move the field toward the definition of new standard models, bringing the practice of visual neuroscience closer to that of established quantitative fields such as Physics. In these fields, there is wide agreement as to what constitutes a standard theory and which results should be the source of surprise.

The review is authored by the speakers and organizers of a mini-symposium at the 2005 Annual Meeting of the Society for Neuroscience. We are all involved in a similar effort: we construct models of neurons and test how accurately they predict the responses to both simple laboratory stimuli and complex stimuli such as those that would be encountered in nature. How accurate are the existing models when held to a rigorous test? By what standards should we judge them? Do they generalize to large classes of stimuli? How should the models be revised?

The review is organized along the lines of the mini-symposium, with each speaker addressing the question “Do we understand visual processing?” at one or more stages of the visual

Received Sept. 2, 2005; revised Oct. 10, 2005; accepted Oct. 11, 2005.

Correspondence should be addressed to Dr. Matteo Carandini, Smith-Kettlewell Eye Research Institute, 2318 Fillmore Street, San Francisco, CA 94115. E-mail: matteo@ski.org.

DOI:10.1523/JNEUROSCI.3726-05.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/2510577-21\$15.00/0

hierarchy. We begin with *Background*, in which we summarize some notions that are at the basis of most functional models in early vision. Demb follows with an evaluation of standard models of the retina (see below, *Understanding the retinal output*). Mante formalizes an extension to the linear model of LGN neurons to account for luminance and gain control adaptation effects (see below, *Understanding LGN responses*). The successes and failures of cortical models are addressed by Tolhurst (see below, *Understanding V1 simple cells*) and Dan (see below, *Understanding V1 complex cells*). Gallant discusses novel model characterization techniques and their degree of success in areas V1 and V4 (see below, *Evaluating what we know about V1 and beyond*). Finally, Olshausen argues that our understanding of V1 is far from complete and proposes future avenues for research (see below, *What we don't know about V1*). In Conclusion, we isolate some of the common ideas and different viewpoints that have emerged from these contributions.

Background

At the basis of most current models of neurons in the early visual system is the concept of linear receptive field. The receptive field is commonly used to describe the properties of an image that modulates the responses of a visual neuron. More formally, the concept of a receptive field is captured in a model that includes a linear filter as its first stage. Filtering involves multiplying the intensities at each local region of an image (the value of each pixel) by the values of a filter and summing the weighted image intensities. A linear filter describes the stimulus selectivity for a neuron: images that resemble the filter produce large responses, whereas images that have only a small resemblance with the filter produce negligible responses. For example, tuning for the spatial frequency of a drifting grating is described by the center-surround organization of filters in the retina and LGN (Fig. 1A) (Enroth-Cugell and Robson, 1966), whereas orientation tuning in V1 is described by filters that are elongated along one spatial axis (Fig. 1B) (Hubel and Wiesel, 1962).

Basic models of neurons at the earliest stages of visual processing (retina, LGN, and V1 simple cells) typically include a single linear filter (Enroth-Cugell and Robson, 1966; Movshon et al., 1978b), whereas models of neurons at later stages of processing (V1 complex cells and beyond) require multiple filters (Fig. 1C) (Movshon et al., 1978b; Adelson and Bergen, 1985; Touryan et al., 2002).

The second stage of these models describes how the filter outputs are transformed into a firing rate response. This transformation typically takes the form of a static nonlinearity (e.g., half-wave rectification), a function that depends only on its instantaneous input. In addition, many models implicitly assume that firing rate is expressed into spike trains via a Poisson process.

Although the receptive field has been described thus far as a set of weights arranged in space (Fig. 1), in reality, the concept of receptive field involves three dimensions: two dimensions of space and the dimension of time. The full spatiotemporal receptive field of a neuron specifies what weight is given to each location in space at each instant in the recent past. When only the temporal evolution of the response is considered for a given spatial position (Fig. 2), the receptive field is commonly referred to as a temporal weighting function.

Whether they are specified in space, time, or jointly in space and time, receptive fields are typically endowed with ON and OFF subfields (Fig. 1, white and black regions). An ON region is one in which a bright light evokes a positive response and a dark light evokes a negative response. An OFF region does the opposite. In

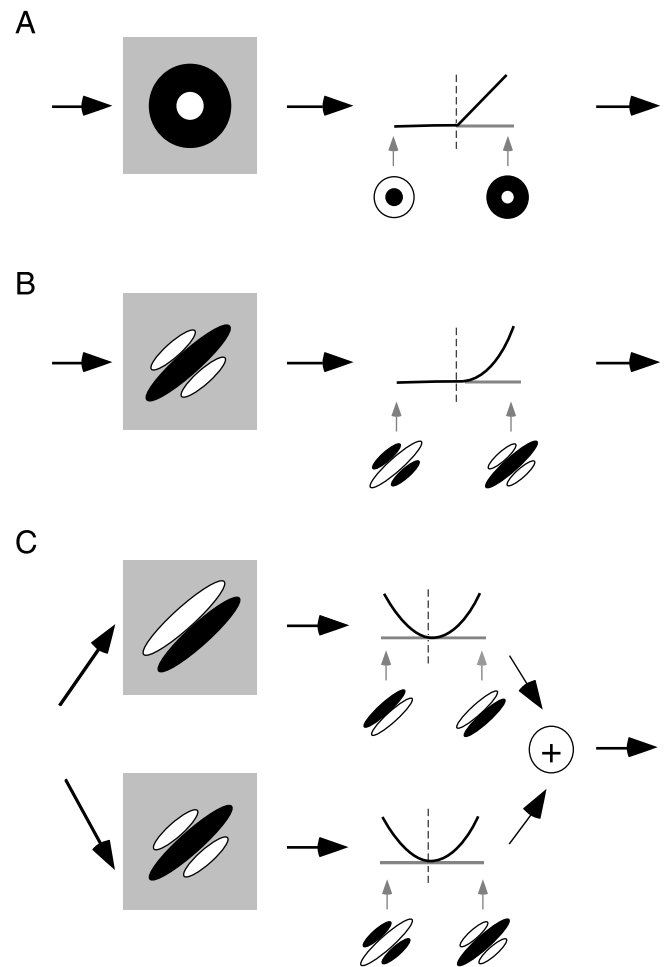


Figure 1. Basic models of neurons involved in early visual processing. In all models, the response of a neuron is described by passing an image through one or more linear filters (by taking the dot product or projection of an image and a filter). The outputs of the linear filters are passed through an instantaneous nonlinear function, plotted here as firing rate on the ordinate and filter output on the abscissa. **A**, Simple model of a retinal ganglion cell or of an LGN relay neuron. The model includes a linear filter (receptive field) with a center-surround organization and a half-wave rectifying nonlinearity. Images that resemble the filter produce large firing rate responses, whereas images that resemble the inverse of the filter or have no similarity with the filter produce no response. **B**, Model of a V1 simple cell as a filter elongated along one axis and a half-wave squaring nonlinearity. As in **A**, only images that resemble the filter produce high firing rate responses. **C**, The energy model of a V1 complex cell. The model includes two phase-shifted linear filters whose outputs are squared before they are summed. In this model, both images that resemble the filters and their inverses produce high firing rates.

the early days, these regions were called “excitatory” and “inhibitory” (Hubel and Wiesel, 1962). However, this name is misleading: their sign has to do with the relative contrast of light, not to the operation of synaptic excitation and inhibition. For instance, an OFF region will deliver substantial excitation in response to a dark stimulus (Hirsch, 2003).

The advantage of assuming an initial linear processing stage is that it enables the experimenter to recover a full model of a neuron within the time constraints of an experiment. Recovering the filter weights involves presenting a sufficiently rich stimulus set to the cell (e.g., white noise, flashed gratings, or natural images) and correlating the response of the neuron with the pixel intensities in the images that immediately preceded spikes. For neurons early in the visual system, a single linear filter is often extracted by presenting a random noise stimulus and computing the mean pixel intensity before each spike, the spike-triggered average

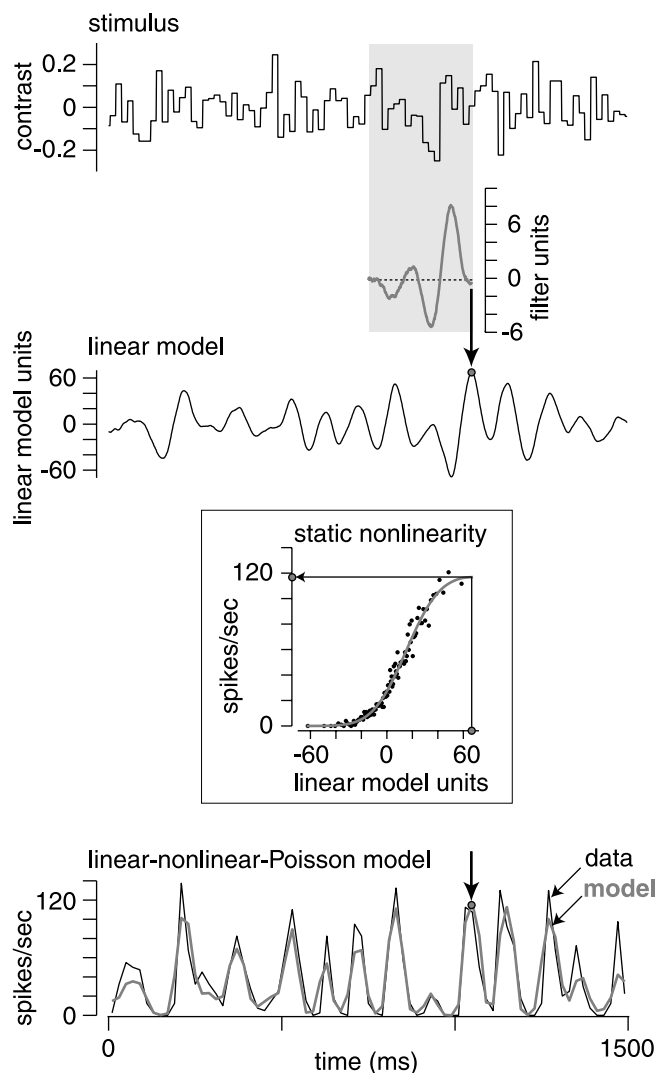


Figure 2. The LNP model of the spike response in a retinal ganglion cell. A model of an ON Y-type ganglion cell was generated from 100 s of response to a white-noise stimulus, contrast was 0.1, extracellular recording in *in vitro* guinea pig retina; methods as described by Zaghloul et al. (2005). The model generates a linear filter (weighting function) and a static nonlinearity (Chichilnisky, 2001). To predict the response to a novel dataset, the stimulus is passed through the filter to generate the linear model of the response. The filter is shown at a time when it closely matched the stimulus (gray box), and so the linear model response is large (+63 in linear model units; gray circle). The linear model is translated to a spike rate using a static nonlinearity that works like a “lookup table” (shown in box). The point in the linear model at +63 is translated to a spike rate of 117 spikes/s. The bottom trace shows the spike rate (black line) to 1.5 s of the novel stimulus (of 2.7 s total). The test stimulus was repeated 20 times, and the data were averaged and binned (bin, 20 ms). The gray line shows the output of the LNP model (bin, 20 ms). The gray circle shows the LNP model value at 117 spikes/s at the same time shown above for the linear model. The r^2 between the data and model was 0.81.

(Chichilnisky, 2001). Similar approaches are followed in the sections below on the retina, lateral geniculate nucleus, and V1 simple cells. At later stages of visual processing, the responses of multiple linear filters can be accounted for by looking at the higher-order correlations between random stimuli and the response of a neuron (Simoncelli et al., 2004). This is the approach followed below in *Understanding V1 complex cells*. Novel nonlinear mapping techniques provide a bridge between these approaches (see below, *Evaluating what we know about V1 and beyond*).

Understanding the retinal output

The retina contains a complex network of cells, divided into an estimated ~60–80 cell types: 3–4 photoreceptors, ~40–50 interneurons, and ~15–20 ganglion cells, whose spike trains transmit visual information to the rest of the brain (Masland, 2001; Sterling and Demb, 2004; Wässle, 2004). No predictive model will suffice for all types of ganglion cell, because some cells have “conventional” center-surround receptive fields (Fig. 1A) (Kuffler, 1953; Enroth-Cugell and Robson, 1966), whereas others have specialized properties, including direction selectivity and intrinsic photosensitivity (Berson, 2003; Taylor and Vaney, 2003; Dacey et al., 2005). As a starting point, predictive models have focused on four ganglion cell types: the ON- and OFF-center versions of sustained (X/parvocellular type) and transient (Y/magnocellular type) cells. These four cell types express relatively simple receptive fields, they project via the LGN to visual cortex, and they can, with some caveats, be modeled in a relatively straightforward way. For the purpose of the predictive models in question, we could ignore all of the complexity of retinal circuitry; the goal is simply to achieve a thorough understanding of how light at the cornea corresponds to spiking responses in the ganglion cell.

Perhaps surprisingly, most retinal studies have not attempted to “go all the way” and predict responses to natural movies, but rather they have focused on a simple dynamic laboratory stimulus: white noise. A white-noise stimulus is created by drawing intensity values from a Gaussian distribution, defined by a mean and an SD of intensity, every ~10–20 ms (Fig. 2). White noise contains approximately equal energy over a range of temporal frequencies (Zaghloul et al., 2005). The relatively flat temporal frequency spectrum is a nice feature for characterizing the receptive field, but this flat spectrum differs markedly from natural scenes, in which there is decreasing stimulus energy at higher temporal frequencies (Simoncelli and Olshausen, 2001). Nevertheless, the response to white noise presents a serious challenge for predictive models and reveals several important nonlinearities.

To take an example, we could perform a simple experiment in which we stimulate a cell with a spot of light over the receptive field center and modulate the spot intensity with white noise (Zaghloul et al., 2005). In this case, we build a model of the temporal response of the cell only [although this approach can easily be extended to model the full spatiotemporal-chromatic response (Chichilnisky, 2001)]. The first step is to build a linear model of the response of the cell. To do so, we cross-correlate the spike response with the white-noise stimulus (Sakai and Naka, 1995; Chichilnisky, 2001). The result is a linear filter that represents the weighting function of the cell (see above, *Background*). Then, at any instant in time, we can generate the linear response by multiplying the stimulus by the temporal weighting function, pointwise, and summing the result (Fig. 2). To generate the linear response at the next moment, we advance the temporal weighting function in time and repeat the process. Under certain conditions, the linear model alone predicts the cone photoreceptor response to a white-noise stimulus (Rieke, 2001; Baccus and Meister, 2002). However, the linear model fails for ganglion cell responses because of several nonlinearities (Shapley and Victor, 1978; Victor, 1987; Chichilnisky, 2001; Kim and Rieke, 2001; Baccus and Meister, 2002; Zaghloul et al., 2003, 2005).

One major nonlinearity is the spike threshold. Resting discharge of ganglion cells can be as low as 0 spikes/s or as high as 80 spikes/s, but a value of 10–20 spikes/s is common (Kuffler et al., 1957; Troy and Robson, 1992; Passaglia et al., 2001). A non-

optimal stimulus will reduce the firing rate, but firing rates cannot go negative, and so there is a point at which spiking responses are “clipped.” Furthermore, an optimal stimulus will increase the spike rate, but spike rates cannot be infinitely high. In a 10 ms period, a cell could fire at most approximately four spikes (or 400 spikes/s) because of the ~ 1 –2 ms refractory period after each spike. Thus, the clipping (“rectification”) and the maximum rate (“saturation”) represent two notable nonlinearities. These nonlinearities can, to some degree, be modeled as “static,” meaning that the linear response can be passed through an input–output function that is invariant over time (Fig. 2) (Chichilnisky, 2001). The combination of a linear filter and a static nonlinearity creates the linear–nonlinear (LN) model of spiking (Figs. 1A, 2). This model predicts the spike rate but not actual spike times; spiking is modeled as a Poisson process, defined by a rate (with equal mean and variance), but spike times are otherwise random. Thus, the model can most properly be termed the linear–nonlinear–Poisson (LNP) model of spiking (Paninski et al., 2004).

Despite its simplicity, the LNP model works rather well at predicting spike rates. In practice, one can generate the linear prediction of the response using the method described above. One can estimate the static nonlinear function by plotting the linear prediction of the response versus the actual response and fitting a smooth function (Chichilnisky, 2001). One way to test the model is to build the linear and nonlinear stages based on one dataset and then test how well the model predicts the response to a novel test stimulus (with the same contrast and mean luminance as the stimulus used to generate the model). On such tests, the LNP model predicts the new dataset nearly as well as does a maximum likelihood “gold standard” (Chichilnisky, 2001; Kim and Rieke, 2001; Zghloul et al., 2003). Another measure is the amount of variance captured by the model (r^2). In Figure 2, the LNP model captured 81% of the variance in the spike response. A similar LN model works equally well on subthreshold membrane voltage or current responses (Kim and Rieke, 2001; Rieke, 2001; Baccus and Meister, 2002; Zghloul et al., 2003, 2005).

We could feel rather satisfied by the ability of the LNP model to predict the response to a novel test stimulus. However, model performance would degrade quickly if we changed almost any aspect of the test stimulus. For example, imagine that we changed the contrast (the SD of the Gaussian distribution of luminance values). Increasing contrast reduces the sensitivity of the linear filter (height) and shortens the integration time (width) (Shapley and Victor, 1978; Smirnakis et al., 1997; Benardete and Kaplan, 1999; Chander and Chichilnisky, 2001; Kim and Rieke, 2001; Zghloul et al., 2005). Thus, to model the response at the new contrast, we would need to use a new filter. However, in many cases, we can model the response at the new contrast with the same nonlinear function as before (Chander and Chichilnisky, 2001; Kim and Rieke, 2001; Zghloul et al., 2005). Still, to predict the response to multiple contrasts, we would need to know the linear filter for each contrast.

Even if we knew the linear filter for all contrast levels, we would have another problem. Each of our linear filters was calculated using a white-noise stimulus with a contrast level that remained constant during the filter measurement. As soon as we move to a natural stimulus, we can expect that the contrast would change continuously, and so we would need to know how the linear filter changed dynamically over time with the contrast level. There is some evidence that the filter changes rapidly after a change in contrast, in ~ 10 –100 ms (Victor, 1987; Baccus and Meister, 2002); however, other measures suggest a slower change over seconds (Kim and Rieke, 2001). Furthermore, there

are cases in which switching contrast to a new level changes not only the filter but also the static nonlinearity (Baccus and Meister, 2002). This introduces a complication for the LNP model because, even if the LNP model is useful at a given, steady contrast level, we must consider that both the linear and nonlinear stages would change dynamically as contrast varied over time in a natural movie.

The above example considers the response to a luminance modulation in time, a one-dimensional problem, but of course a natural movie varies over time in two dimensions of space (plus there is the issue of color). When we consider space, two complications arise. First, transient (Y-type) ganglion cells combine subregions of their receptive field nonlinearly, apparently because of nonlinearities at the output of presynaptic bipolar cells (Demb et al., 2001). Furthermore, there are nonlinear signals passed across the retina from outside the classical (center–surround) receptive field that are not captured by the LNP model (Demb et al., 1999; Roska and Werblin, 2001; Olveczky et al., 2003). Some models have characterized these nonlinear influences using quantitative approaches (Shapley and Victor, 1978; Victor, 1979). However, it is not clear at present how well these models would predict responses to natural stimuli. Furthermore, certain ganglion cells adapt to the pattern of light over space or time, such that the linear filter becomes less sensitive to the most predictable features of the stimulus (Hosoya et al., 2005). For example, this type of adaptation would increase sensitivity to horizontal features after prolonged exposure to vertical features. This pattern adaptation will need to be considered in future predictive models.

One direction to push the LNP model is to generate a more realistic pattern of spiking than Poisson output. In fact, ganglion cells, unlike cortical cells, fire spikes much more reliably than a Poisson process (Berry et al., 1997; Reich et al., 1997; Kara et al., 2000; Demb et al., 2004; Uzzell and Chichilnisky, 2004). For example, a stimulus that evokes, on average, a burst of nine spikes will show an SD (across repeated trials) of approximately one spike rather than the Poisson value of three spikes (i.e., variance of 9, equal to the mean). One recent approach used a novel method for fitting a model that includes a linear filter followed by an integrate-and-fire spike generator (Paninski et al., 2004). To model realistic patterns of spiking, the spike generator includes a recovery function, after each spike, mimicking a refractory period (Keat et al., 2001; Paninski et al., 2004). One result of this approach is that the apparent contrast-dependent change in the linear filter width as measured by the LNP model may be an artifact related to the refractory period in the data (Pillow and Simoncelli, 2003). However, intracellular studies, which measure the continuous subthreshold potential, suggest that some amount of the contrast-dependent change in filter width may be real (Kim and Rieke, 2001; Zghloul et al., 2005).

Even given all of the above complications, it is surprising that more retinal studies have not attempted to predict the response to a natural movie. One group tested their model on a full-field stimulus that was modulated by a natural sequence of light fluctuation, and the model did a reasonable job (van Hateren et al., 2002). The model was based on a linear filter approach and included feedback gain controls, to account for adaptation to the mean intensity and contrast, and a rectifying nonlinearity to model the spike threshold. In fact, many “bursts” of spiking evoked during the stimulus were captured by the model, although there was clearly room for improvement. Furthermore, the study did not test the predictive power of the model on novel datasets. Still, the results were generally encouraging.

What are the next steps for predictive modeling in the retina? Clearly, there are many questions left unanswered by the LNP model. A major question is how we can predict the linear filter at any instant in time, given the previous statistics of the stimulus. A working hypothesis suggests that the retina adapts separately to contrast (“contrast adaptation”) and the mean intensity (“light adaptation”). So one advance would be to further understand the rules by which the previous mean intensity and contrast influence the filter (see below, *Understanding LGN responses*). However, this hypothesis suggests that the mean and contrast are the only relevant parameters and that these parameters control filter adaptation independently; both of these assumptions require additional validation. Also, this theory ignores possible adaptation to higher-order stimulus statistics (Hosoya et al., 2005). Furthermore, once we get past photoreception and into the retinal circuit, cells no longer adapt to light statistics; rather, they adapt based on changes in neurotransmitter release over time as well as intrinsic cellular properties. Thus, it will be important to further understand cellular mechanisms for adaptation. Intracellular recordings suggested that contrast adaptation occurs partly at the level of synaptic input and partly at the level of ganglion cell spike generation, suggesting an adaptive mechanism intrinsic to the ganglion cell (Kim and Rieke, 2001, 2003; Zaghloul et al., 2005). Further understanding cellular mechanisms for adaptation could provide key insights into the optimal architecture of a predictive model. In other words, we should amend the statement at the beginning of this section about ignoring retinal circuitry: knowledge of the circuitry could indeed guide the development of an appropriate predictive model.

In summary, the response to a dynamic laboratory stimulus, white noise, can be predicted fairly well using a simple linear model followed by a static nonlinearity. Rapid changes in stimulus statistics, as would occur in a natural movie, requires additional understanding of the rules by which the linear and nonlinear stages adapt over time. Furthermore, there are advances to be made in modeling spike times, as opposed to rates, and we need to further understand the multiple ganglion cell types beyond the four types considered here. To many in the field of visual neuroscience, predicting responses in the retina seems much simpler than predicting responses in an extrastriate area, such as V4, and there is clearly truth to this notion. Nevertheless, there is still a ways to go before we can predict retinal responses to an arbitrary stimulus.

Understanding LGN responses

The LGN occupies a strategic position, a strait through which most retinal signals must pass to reach visual cortex. The strongest retinal input to the LGN originates from ganglion cells of the X/parvocellular type and of the Y/magnocellular type. These two cell types have been studied extensively (see above, *Understanding the retinal output*) and together constitute ~50% of ganglion cells in the cat and ~80% in primates (Rodieck et al., 1993; Masland, 2001; Wässle, 2004). Additional input to LGN relay cells originates from other geniculate neurons, from subcortical structures, and from cortex (Guillery and Sherman, 2002).

It would be highly desirable to obtain a complete description of how LGN neurons respond to visual stimuli. Such a description would summarize the computations performed by the retinal and thalamic circuitry and amount to a full understanding of the visual inputs received by primary visual cortex.

The main determinant of the responses of LGN neurons is the linear receptive field, whose broad attributes are similar to those of the afferent retinal ganglion cells. The receptive field is com-

posed of a center and of a larger surround, whose responses interact subtractively (Fig. 1A). Both center and surround have a biphasic temporal weighting function (Fig. 2), i.e., they weigh contributions from the recent and less recent past with opposite polarity (Cai et al., 1997; Reid et al., 1997). The linear receptive field accurately predicts the basic selectivity of LGN neuron measured with gratings. For instance, the spatial profile of the receptive field predicts the selectivity for spatial frequency (Kaplan et al., 1979; So and Shapley, 1981; Shapley and Lennie, 1985), whereas the temporal weighting function predicts the selectivity for temporal frequency (Saul and Humphrey, 1990; Kremers et al., 1997; Benardete and Kaplan, 1999). The linear receptive field does not describe only responses to simple laboratory stimuli but also captures the basic features in the responses to complex video sequences (Dan et al., 1996).

The shape of the temporal weighting function of LGN neurons depends on two strong nonlinear adaptive mechanisms that originate in retina: luminance gain control and contrast gain control (see above, *Understanding the retinal output*) (Shapley and Enroth-Cugell, 1984). These gain control mechanisms affect the height (i.e., the gain) and width (i.e., the integration time) of the temporal weighting function. Luminance gain control (also known as light adaptation) occurs primarily in retina. It matches the limited dynamic range of neurons to the locally prevalent luminance (light intensity). Gain and integration time are reduced for locations of the visual field where mean luminance is high and increased where mean luminance is low (Dawis et al., 1984; Rodieck, 1998). Contrast gain control begins in retina (Shapley and Enroth-Cugell, 1984; Victor, 1987; Baccus and Meister, 2002) and is strengthened at subsequent stages of the visual pathway (Kaplan et al., 1987; Sclar et al., 1990). It regulates gain and integration time on the basis of the locally prevalent root-mean-square contrast, the SD of the stimulus luminance divided by the mean luminance. Gain and integration time are reduced for locations of the visual field in which contrast is high and increased in which contrast is low.

These gain control mechanisms dampen the impact of sudden changes in the mean luminance or contrast of a scene such as those brought about by eye movements. This effect is illustrated in Figure 3, A and B, by the responses of an LGN neuron in an anesthetized, paralyzed cat (Mante et al., 2005). Stimuli were drifting gratings of optimal spatial frequency. Several seconds after the onset of a grating, either mean luminance (at constant contrast) or contrast (at constant mean luminance) was suddenly increased. LGN responses are barely affected by the change in luminance (Fig. 3A) and only weakly affected by the change in contrast (Fig. 3B). Indeed, consider the responses predicted at high luminance from the linear receptive field measured at low luminance (Fig. 3A, red curves) and the responses predicted at high contrast from the linear receptive field measured at low contrast (Fig. 3B, red curves). The linear predictions are larger and slower than the measured responses, indicating that gain and integration time are reduced when luminance or contrast are increased. This reduction in gain and integration time is completed within a cycle of the drifting grating, demonstrating that the gain control mechanisms operate in <100 ms (Enroth-Cugell and Shapley, 1973a; Saito and Fukada, 1986; Victor, 1987; Lankheet et al., 1993a; Yeh et al., 1996; Baccus and Meister, 2002; Lee et al., 2003; Mante et al., 2005). This fast timescale suggests that gain control mechanisms reduce the impact of eye movements, which place the receptive field of neurons in the early visual system over regions of widely different mean luminance and contrast (Mante et al., 2005).

Although the gain control mechanisms are likely to play a major role during natural vision, most efforts to predict the responses of LGN neurons to natural stimuli have been limited to assuming a fixed linear receptive field (Dan et al., 1996) and have omitted the effects of gain control. There are several reasons for this omission. First, existing models of gain control are limited in scope: they operate only on simplified stimuli such as gratings, and they lack a definition of luminance and contrast that applies to arbitrary stimuli. Second, with few exceptions (Troy and Enroth-Cugell, 1993), luminance gain control and contrast gain control were typically only studied in isolation. During natural vision, however, luminance and contrast vary independently of each other (Mante et al., 2005). Thus, a general model of gain control should predict the shape of the temporal weighting function at every possible combination of luminance and contrast.

Nonetheless, many of the components needed to build a general model of gain control have already been described individually. For instance, studies have separately modeled the effects of luminance gain control (Fuortes and Hodgkin, 1964; Baylor et al., 1974; Brodie et al., 1978; Shapley and Enroth-Cugell, 1984; Purpura et al., 1990) and of contrast gain control (Shapley and Victor, 1981; Victor, 1987; Carandini et al., 1997; Benardete and Kaplan, 1999) on the temporal weighting function of neurons in the early visual system. Many models share the same simple design: the temporal weighting function is obtained by convolving a fixed temporal weighting function with a variable filter, whose parameters depend on luminance or contrast. This design can be easily extended to predict the temporal weighting function at any combination of luminance and contrast. Indeed, a recent study of LGN responses demonstrated that the effects of luminance gain control and contrast gain control are independent of each other (Mante et al., 2005). Thus, the temporal weighting function can be described by convolving the fixed weighting function with two variable filters, one that depends on luminance and the other that depends on contrast.

These studies provide a number of clues about how retinal or LGN neurons compute the luminance and contrast of an arbitrary stimulus. Luminance and contrast are computed not only very rapidly (Fig. 3*A,B*) but also very locally. Luminance gain control is driven by the average light intensity falling onto a region that is not larger than the surround of the linear receptive field (Cleland and Enroth-Cugell, 1968; Enroth-Cugell and Shapley, 1973b; Enroth-Cugell et al., 1975; Cleland and Freeman, 1988; Lankheet et al., 1993b). Similarly, contrast gain control is driven only by stimuli lying within the linear receptive field (Solomon et al., 2002; Bonin et al., 2005). More precisely, contrast seems to be computed from the integrated responses of a pool of small, nonlinear subunits coextensive with the linear receptive field (Shapley and Victor, 1979; Enroth-Cugell and Jakiela, 1980; Bonin et al., 2005).

These insights on gain control can be used to build a nonlinear

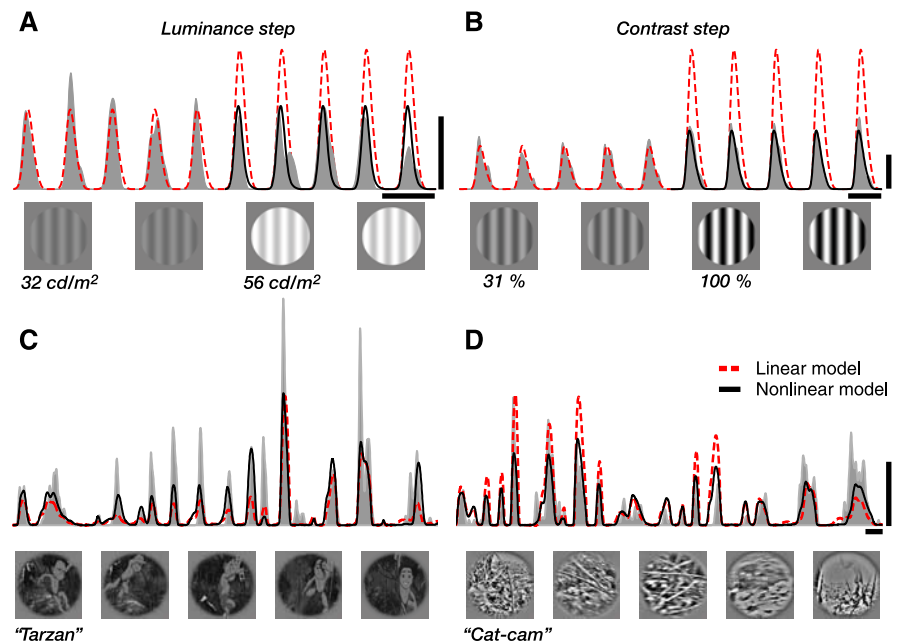


Figure 3. Predicting responses of LGN neurons to complex video sequences. **A**, Firing rate responses of an LGN neuron to a drifting grating whose mean luminance is suddenly increased from 32 to 56 cd/m^2 (while contrast is kept constant). Red dashed traces indicate the prediction of the linear receptive field fitted to the response before the luminance step, and black solid traces indicate the average response after the step. **B**, Same, for a stimulus whose contrast suddenly steps from 31 to 100% (while mean luminance is kept constant). **C**, Responses of an LGN neuron to a sequence from Walt Disney's *Tarzan*. Red dashed traces indicate the prediction of the linear receptive field alone (measured at optimal luminance and contrast). Black solid traces indicate the prediction of a nonlinear model. In the nonlinear model, the gain and integration time of the receptive field are regulated by luminance gain control and contrast gain control. **D**, Same, for responses to a *Cat-cam* movie (Kayser et al., 2003; Betsch et al., 2004). In all panels, calibration is 100 ms and 100 spikes/s, and gray histograms are firing rates obtained by convolving the spike trains with a Gaussian window of width 5 ms (SD). **A** and **B** are modified from Mante et al. (2005b). **C** and **D** are modified from Mante (2005).

model of LGN responses that is general enough to predict the responses to arbitrary stimuli (Bonin et al., 2005; Mante, 2005). This model predicts a number of nonlinear phenomena in the responses to simple stimuli, none of which would be explained by the linear receptive field alone. (1) Response amplitude is independent of mean luminance at low temporal frequencies, although it is approximately proportional to mean luminance at high frequencies (Shapley and Enroth-Cugell, 1984; Purpura et al., 1990). (2) Response amplitude saturates with contrast. As contrast is increased, the gain is decreased, although not so much as to make responses independent of contrast ["contrast saturation" (Derrington and Lennie, 1984; Cheng et al., 1995)]. (3) Responses are selective for stimulus size, being maximal for stimuli of intermediate size and being suppressed by larger stimuli ["size tuning" (Jones et al., 2000; Solomon et al., 2002; Ozeki et al., 2004)]. For large stimuli, an increase in size adds only little excitatory drive to the responses, although it strongly reduces gain by recruiting more of the subunits driving contrast gain control. (4) The strength of contrast saturation and size tuning depends on the temporal frequency of the stimulus. Both are strong at low temporal frequencies but absent at high temporal frequencies (Shapley and Victor, 1978; Sclar, 1987; Mante et al., 2004). (5) The response to a test stimulus is reduced by superposition of a mask stimulus ["masking" (Freeman et al., 2002; Bonin et al., 2005)].

The nonlinear model predicts the responses to complex, natural stimuli better than the linear receptive field alone (Mante, 2005). For example, the gray histograms in Figure 3, **C** and **D**, represent the firing rate of an LGN neuron in response to two

complex stimuli: movies taken from the head of a cat roaming through a forest [*Cat-cam* (Kayser et al., 2003; Betsch et al., 2004)] and segments of cartoons (Walt Disney's *Tarzan*). The linear receptive field, which has the same temporal weighting function throughout the movies, predicts the basic features of the response but not the details (red curves). In particular, it captures the timing of the responses but not their amplitude. The predictions of the nonlinear model, in which gain and integration time are adjusted dynamically, are more accurate than those of the linear receptive field (black curves). Because of luminance gain control, the predictions of the nonlinear model tend to be higher than those of the linear model during the dark *Tarzan* movie and lower during the bright *Cat-cam* movie. Because of contrast gain control, the two models make different predictions about the relative magnitude of the responses within a movie.

The comparison between the simpler linear model and the more complex nonlinear model is fair because the models were given the same number of free parameters (two: spontaneous firing rate and maximal firing rate). The remaining parameters were estimated from the responses to gratings and then fixed in the predictions of the responses to complex stimuli. Fixing the parameters in advance is necessary to compare the predictions on an equal footing: the more complex nonlinear model does not necessarily have to predict the data better than the simpler linear model. In fact, given the complexity of the nonlinear model, it would have been difficult to estimate its parameters directly from the responses to complex stimuli. This approach might be useful to characterize also later stages of visual processing, in which neurons exhibit progressively more nonlinear properties.

Even the nonlinear model, however, fails to capture some features in the responses. In particular, the measured responses tend to be more transient than the predicted responses. One factor contributing to the transient responses could be the mechanisms generating bursts of action potentials: after a hyperpolarization lasting 100 ms or longer, LGN neurons are likely to respond to a depolarization with a burst of action potentials that is not predicted by simple rectification of the membrane potential (for review, see Sherman, 2001). Bursts are a prominent feature of LGN responses in anesthetized or sleeping animals but less so in awake animals (Guido and Weyand, 1995; Ramcharan et al., 2005). Another factor contributing to the transient responses lies in the spike generation mechanisms: firing rates are more transient than predicted by a simple rectification of the membrane potential (see above, *Understanding the retinal output*). Both mechanisms could be easily incorporated into the nonlinear model (Mukherjee and Kaplan, 1995; Smith et al., 2000; Keat et al., 2001; Lesica and Stanley, 2004; Paninski et al., 2004). Finally, the nonlinear model might be easily extended to capture also the nonlinear spatial properties of Y-cells. In fact, at least in retina, the output of Y-cells can be thought of as the sum of a pool of nonlinear subunits, similar to the one driving contrast gain control (Hochstein and Shapley, 1976; Victor and Shapley, 1979; Enroth-Cugell and Freeman, 1987; Demb et al., 2001).

In summary, there is now a fairly good understanding of the linear and nonlinear components required to model responses of the broad majority of LGN neurons. Many nonlinear properties of LGN neurons can be captured by a single model that is general enough to operate on arbitrary stimuli that vary in both space and time. This model will be a useful tool to explore the effects of gain control during natural vision. Once extended with bursting and spiking mechanisms, it promises to provide a tractable description of the responses of LGN neurons and thus, of the input to primary visual cortex.

Understanding V1 simple cells

Simple-cell receptive fields were first described in area V1 of the cat by Hubel and Wiesel (1959), who defined them as follows: "... these fields were termed 'simple' because like retinal and geniculate fields (1) they were subdivided into distinct excitatory and inhibitory regions; (2) there was summation within the separate excitatory and inhibitory parts; (3) there was antagonism between excitatory and inhibitory regions; and (4) it was possible to predict responses to stationary or moving spots of various shapes from a map of the excitatory and inhibitory areas." (Hubel and Wiesel, 1962).

If a neuron failed any part of this four-part definition (particularly point 1), then it would be termed a "complex cell." These definitions were qualitative; many subsequent studies have enquired whether successful quantitative definitions are possible.

Point 4 is crucial to the definition: can a straightforward receptive-field map predict how the neuron responds to other visual stimuli? We must first acknowledge that predicting responses to time-varying stimuli (e.g., moving ones) requires knowledge of the time courses of responses in different parts of the receptive field. As explained above in *Background*, a static receptive field should be replaced by a spatiotemporal receptive field map (McLean and Palmer, 1989; Reid et al., 1991; DeAngelis et al., 1993a), which documents differences in response time course (impulse or step response) in different parts of the field (Movshon et al., 1978a; Dean and Tolhurst, 1986). The essence of prediction is the same, but, strictly, the field and stimuli should be considered as functions of time as well as functions of space.

Those who follow Hubel and Wiesel's definitions of simple and complex cells generally find the two classes of neuron in approximately equal numbers in V1. The clear definition of a simple cell has been massively influential in visual science because it offers the promise that, from relatively simple experiments, we may understand how approximately half of the neurons in V1 would respond in more complex situations, such as viewing of natural scenes. The definition says essentially that spatiotemporal summation in simple cells is linear: only very simple arithmetic is needed to calculate how a given simple cell will respond to some arbitrary stimulus. Such is the starting point for modeling human psychophysical experiments (Watson, 1987) or for hypothesizing how natural information may be most efficiently coded in V1 (Willmore and Tolhurst, 2001). The simple-cell definition offers so much that we are reluctant to ask whether it really works. This section asks whether simple experiments on simple cells really do allow quantitative predictions about the responses to other, more complicated stimuli.

Movshon et al. (1978a,b) examined the linearity of spatial summation in simple and complex cells in cat V1, followed by Andrews and Pollen (1979). These studies compared spatial receptive-field maps with the tuning for sinusoidal gratings and found that important aspects of summation in simple cells were, indeed, linear when tested quantitatively. Later studies have convincingly shown that the spatiotemporal receptive field of a simple cell precisely predicts the optimal orientation, spatial frequency, and temporal frequency of sinusoidal gratings of the neuron (Movshon et al., 1978a; Jones and Palmer, 1987; Tadmor and Tolhurst, 1989; DeAngelis et al., 1993b; Gardner et al., 1999). However, even the 1970s experiments noted nonlinearities in simple-cell behavior. A saving device was to consider the simple cell as a black box with two stages: a linearly summing stage, followed by one or more nonlinearities that do not affect the underlying initial linear sum (Fig. 1B).

Although the spatiotemporal receptive field of a cell predicts

the cell's optimal stimulus well, it poorly predicts the relative magnitude of responses to nonoptimal stimuli (e.g., it overestimate the bandwidths of orientation and frequency tuning curves) (Jones and Palmer, 1987; Tadmor and Tolhurst, 1989; DeAngelis et al., 1993b; Gardner et al., 1999). Linear models also fail to explain the relative magnitudes of response in the two directions of movement orthogonal to the preferred orientation of the neuron (Albrecht and Geisler, 1991; Reid et al., 1991). These failures of the linear model have an easy rationalization in the two-stage model (Fig. 1*B*). Most experiments record neuronal response extracellularly as trains of action potentials, but the membrane potential changes of V1 neurons must exceed a threshold before spiking activity is evident (Carandini and Ferster, 2000). Many failures of the simple linear model can be accounted for arithmetically, by supposing that a simple linear sum is transformed at the output of the neuron by passage through a nonlinear transducer function; this may simply have a threshold nonlinearity or might be a sigmoidal function of stimulus contrast (Schumer and Movshon, 1984; Tolhurst and Dean, 1987, 1991; Tadmor and Tolhurst, 1989; Albrecht and Geisler, 1991; DeAngelis et al., 1993b). Indeed, intracellular recordings in simple cells (which presumably show the black box before any nonlinear output transform) do suggest that the strength of directional selectivity and the orientation tuning bandwidth can be described by a linear model in the first stage of the black box (Jagadeesh et al., 1993; Lampl et al., 2001).

Post hoc application of a nonlinear transducer to the linear prediction may work arithmetically, but there are inconsistencies between experiments (Tolhurst and Heeger, 1997b). Furthermore, there are other nonlinear behaviors that cannot be explained in such a way. Notable nonlinearities (shared with complex cells) are response saturation at high contrasts (Albrecht and Hamilton, 1982), and "nonspecific suppression" (Bonds, 1989; DeAngelis et al., 1992; Tolhurst and Heeger, 1997b) in which the response of a simple cell to its optimal stimulus is suppressed by simultaneously presenting stimuli that evoke no overt response when presented alone. Heeger (1992a,b) proposed a neuronal circuit that embraces these and many other nonlinear behaviors of simple cells: essentially, each simple cell performs a first-stage linear sum of its spatiotemporal inputs, "half-squares" that linear sum giving an energy response (half-squaring achieves much the same as a threshold nonlinearity), and is then subject to divisive inhibition from all other neurons whose receptive fields cover the same part of visual field. The divisive inhibition gives rise to "contrast normalization." Application of this model (Heeger, 1993; Tolhurst and Heeger, 1997a) resolves subtle failures (Reid et al., 1991; Tolhurst and Dean, 1991) in the predictions of the relative magnitudes of response to moving and stationary-modulated gratings, which cannot be resolved by simply running a linear response sum through a nonlinear output transducer.

Elaborations of the contrast-normalization model (Carandini and Heeger, 1994; Carandini et al., 1997) have embraced additional nonlinear behaviors (previously unaccounted), such as "phase advance" at high contrasts (Dean and Tolhurst, 1986). The contrast-normalization model has been influential in psychophysical modeling (Watson and Solomon, 1997) as well as in understanding the details of simple-cell (and complex-cell) responses; it is ironic that its proponents now suggest a very different neurophysiological mechanism (Carandini et al., 2002; Freeman et al., 2002), although the arithmetic remains more or less the same.

Nonspecific suppression results from stimuli within the receptive field. There is another nonlinearity, sometimes confused

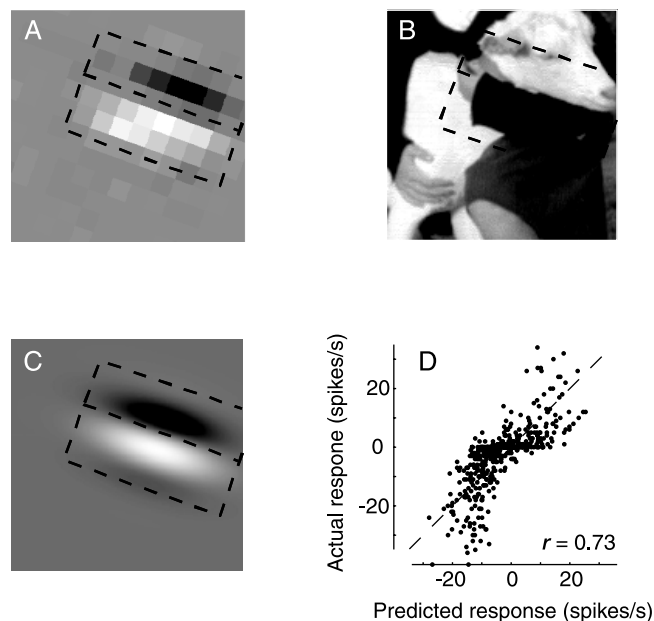


Figure 4. Predicting responses of V1 simple cells to complex images. *A*, The receptive field of a ferret simple cell (see Fig. 1*B*). The dashed lines outline the field for comparison in *B* and *C*. *B*, The digitized photograph that evoked the largest OFF response in this simple cell. *C*, A Gabor patch that approximately matches the disposition of the actual receptive field in *A*. *D*, The ordinate plots the actual responses of the simple cell to 500 different photographs (average of 10 presentations each, measured in 100 ms bins). Positive values are ON responses: spikes generated during the 100 ms presentation. Negative values are OFF responses: spikes generated on removal of the photograph. The abscissa shows the responses predicted on a totally linear model, which includes solely the receptive field shown in *C* and not the nonlinearity of the output (see Fig. 1*B*).

with it: stimuli outside the "classical receptive field" of a simple cell may also suppress or facilitate its responses to its preferred stimuli, as first described by Blakemore and Tobin (1972) and Maffei and Fiorentini (1976). It is growing clearer that different mechanisms of suppression are involved within the classical receptive field and outside (Sengpiel et al., 1998; Freeman et al., 2002; Li et al., 2005; Sengpiel and Vorobyov, 2005), but we do not yet have simple arithmetic rules to describe these nonlinearities. Suppression or facilitation from outside the classical receptive field may result from local connections within V1 or from feedback from more anterior visual areas, perhaps subserving selective attention or perceptual grouping. There is a large literature on this topic that is beyond the present scope (for review, see Fitzpatrick, 2000; Freeman et al., 2001; Chisum and Fitzpatrick, 2004).

Thus, many nonlinearities of simple-cell behavior are not evident from the receptive-field structure, but they can be accommodated neatly into the two-stage black box: the linearly summing first stage is followed by half-squaring and contrast normalization. The arithmetic is fairly easy and it works well. However, how important are all of these nonlinearities in the overall behavior of the neurons? Smyth et al. (2003) examined how simple cells in anesthetized, paralyzed ferret V1 respond to 100 ms flashes of digitized photographs of natural scenes to understand how neurons might respond under natural vision. Others have also sought to understand how the receptive field structure or grating responses of simple cells relate to their responses to complex, natural scene stimuli (Ringach et al., 2002; Vinje and Gallant, 2002; Weliky et al., 2003; David et al., 2004). Figure 4*A* shows the spatial receptive field of one simple cell recorded by

Smyth et al. (2003), mapped with small bright and dark squares. According to point 4 of the simple-cell definition, we expect this neuron to respond particularly well to a bright-dark border, slightly off horizontal and toward the top right of the stimulus area. Figure 4*B* shows the photograph that elicited (by far) the most activity from this neuron. There is, indeed, just the border predicted, although its polarity is reversed compared with the ON and OFF regions of the field: the photograph evoked strong OFF responses. Figure 4*C* shows a Gabor function model of the receptive field (Field and Tolhurst, 1986; Jones and Palmer, 1987), fitted by eye. It was used to estimate how a totally linear field might respond to the 500 photographs presented; no nonlinearities are modeled in the calculation. Figure 4*D* plots the actual response of the simple cell to the photographs against the responses predicted from the stylized linearly summing field. The linear model conveys the gist of the actual responses ($r = 0.73$); there are no astonishing outliers. In truth, the simple cell of Figure 4 is the one whose responses to natural scenes were best predicted by linear modeling (Smyth et al., 2003). However, the results for this and other simple cells suggest that, although output nonlinearities may reduce response magnitudes below the linear prediction, there is little evidence here that nonlinear effects could fundamentally alter the basic “trigger features” for activating a simple cell.

It is important also to recognize that simple cells are heterogeneous; some simple cells may differ little from complex cells (Dean and Tolhurst, 1983; Mechler and Ringach, 2002). Movshon et al. (1978a) described some “nonlinear simple cells” in which the ON and OFF receptive-field regions overlap; Dean and Tolhurst (1983) found that receptive-field structure was continuously graded from simple cells exactly fitting the Hubel and Wiesel definition, through nonlinear simple cells and “discrete complex cells” to frank complex cells (Priebe et al., 2004; Mata and Ringach, 2005). Indeed, simple and complex cells may not form a dichotomy at all. Of course, this is not to say that all cells are the same; for instance, it is clearly understood that cells at the “simple end” of the continuum are found in different cortical layers than those at the “complex end” (Martinez et al., 2005). Receptive-field mapping techniques typically subtract the responses, say, to dark stimuli from those to bright stimuli so that the receptive field seems to be single valued at each point. The resulting linear receptive field is an incomplete reflection of the overall responses of the neuron. For the varied population of simple cells, it is unclear what proportion of response is dependent only on the idealized linear receptive field and what has been obscured by ignoring the inherent nonlinearities of summation.

The two-stage model of Figure 1*B* is inaccurate; its simplicity and convenience may be misleading. Geniculate inputs are inherently nonlinear and may be subject to depression leading to the nonspecific suppression noted above (Carandini et al., 2002). Nonlinear inputs would result in inherent nonlinearities of simple-cell summation unless, say, there were “push–pull” inhibition (Glezer et al., 1982). The role of such inhibition has been explored further (Palmer and Davis, 1981; Tolhurst and Dean, 1987; Ferster, 1988; Tolhurst and Dean, 1990; Hirsch et al., 1998). In particular, Tolhurst and Dean (1987, 1990) and Atick and Redlich (1990) proposed that breakdown of push–pull inhibition underlies the appearance of nonlinearities of spatial summation even in the first stage of the two-stage model (Fig. 5). Indeed, all that may distinguish many complex cells from simple cells might just be the strength of the inhibitory signals that mask inherently nonlinear summation (Wieland et al., 2001; Mechler and Ringach, 2002).

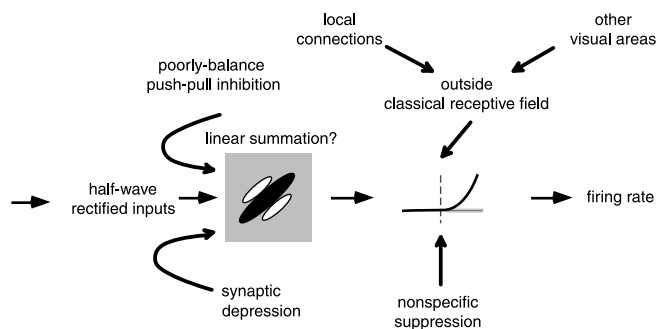


Figure 5. Toward a complete model of V1 simple cells. A popular conception of simple-cell response behavior involves a first stage that shows strictly linear spatiotemporal summation, and a subsequent stage may be subject to a variety of nonlinear phenomena, which do not impinge on the fundamental linearity of the first stage (see Fig. 1*B*). This schematic shows a revision of this convenient model, which includes a number of nonlinear mechanisms. Some of these mechanisms (those depicted as affecting the output nonlinearity) spare the fundamental linearity of summation. The remaining ones, however, cause nonlinear summation, which differs only in degree from the obvious nonlinear behavior of complex cells.

Literal adherence to point 4 of the definition of a simple cell means that any significant failure of prediction would require the neuron to be reclassified as a complex cell. Thus, we would be bound to understand simple cells; any problematic neuron must be a complex cell. Failure of the linear receptive field model is not a problem for neurophysiologists, but it is for those computational modelers who would like all neurons in the visual cortex to be described in a few lines of elegant code with their receptive-field parameters neatly spaced along a theorist's dimensions. We need not confuse failure of the attractive linear receptive field model of the simple cell with failure to understand the visual cortex as a whole (compare with below, *What we don't know about V1*). Many of the bolt-on nonlinearities of simple cell behavior can be parameterized coherently; the failures of push–pull are an irritation to modelers, but there is no need to believe that they portend any dramatic change in neuronal behavior over the linear model, and, most significantly, much progress has been made in understanding how (frankly nonlinear) complex cells respond to naturalistic stimuli (compare with below, *Understanding V1 complex cells*).

Understanding V1 complex cells

Although the orientation and spatial frequency selectivity of each simple cell is directly related to the spatial profile of its receptive field, which consists of elongated ON and OFF subregions (see above, *Understanding V1 simple cells*), for complex cells this relationship is far less obvious. A complex cell usually exhibits mixed ON and OFF responses throughout its receptive field (Hubel and Wiesel, 1962). For example, the response of the cell to a bar stimulus depends on both the orientation and width of the bar in a manner similar to simple cells, but the cell responds indiscriminately to light and dark bars, as long as the bar stands out from the gray background. Such insensitivity to contrast polarity is an important form of nonlinearity that renders spike-triggered average ineffective for measuring complex-cell receptive fields.

Significant progress in understanding complex-cell receptive fields was first made by measuring the nonlinear interactions between a pair of bars at the preferred orientation of the cell (Movshon et al., 1978b; Emerson et al., 1987). These studies have revealed the existence of “subunits” of the complex-cell receptive field, whose spatial structure can predict the frequency tuning of the cell. More recently, an alternative method has been used to

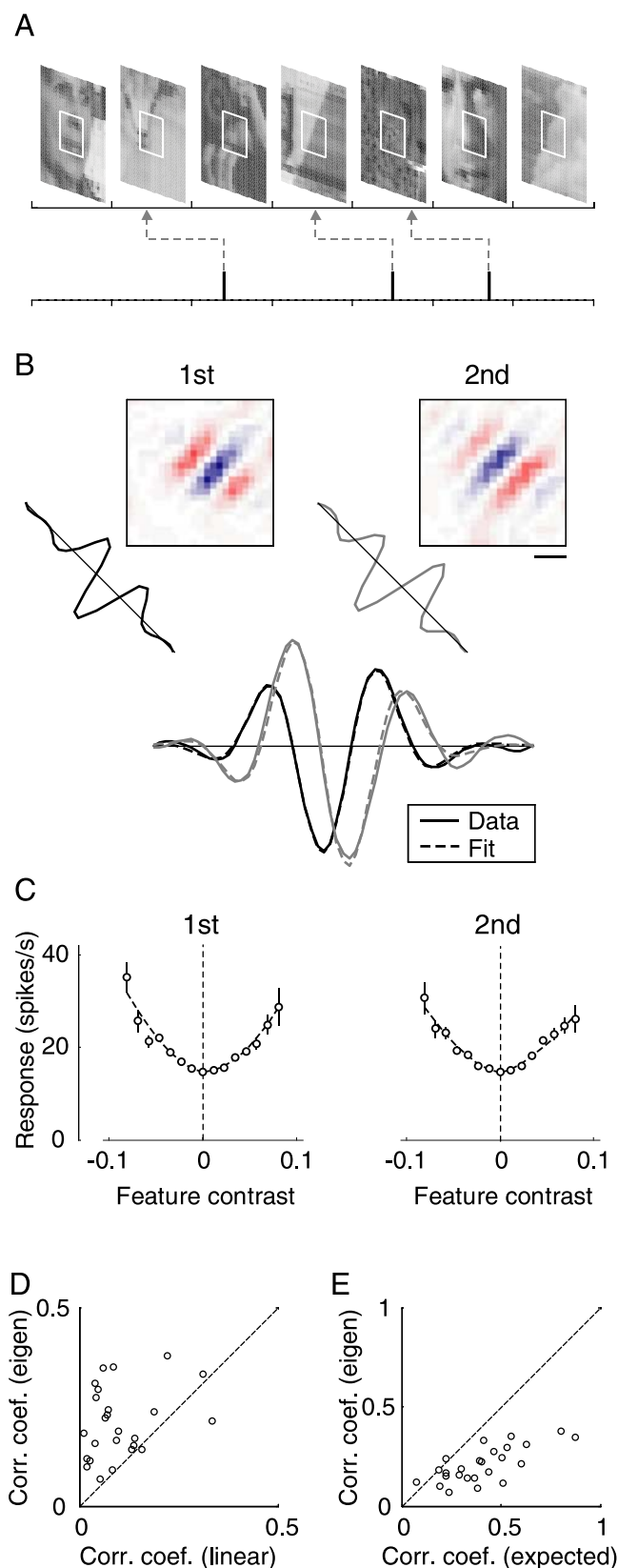


Figure 6. Models of V1 complex cells recovered from a covariance analysis. **A**, Experimental protocol. Top, A segment of the natural image ensemble; white box indicates area shown in experiments. Bottom, Spike train. Spike-triggered ensemble was generated by collecting the image preceding each spike by a single frame (42 ms). **B**, Two significant eigenvectors of a complex cell. Scale bar, 2° . Solid line, Spatial profiles of each eigenvector along the axis perpendicular to the preferred orientation. Dashed line, Gabor fit. The Gabor fits of the two eigenvectors had a phase difference of 85° . **C**, Contrast–response functions of the two eigenvectors.

characterize complex-cell receptive fields that uses large ensembles (tens of thousands) of complex visual stimuli, such as white noise or natural images, and a spike-triggered covariance (STC) analysis to receptive-field estimation.

As a first step in STC, the stimulus preceding each recorded spike is collected to form the spike-triggered stimulus ensemble, just like in spike-triggered average. Then, the covariance matrix, instead of the mean, of this spike-triggered ensemble is computed. Eigenvectors of this matrix with “significant eigenvalues” (those significantly different from the control eigenvalues calculated based on random spike trains) represent visual features that directly affect the neuronal response. This method has been used effectively to analyze the nonlinear response properties of fly visual neurons (De Ruyter Van Steveninck and Bialek, 1988; Brenner et al., 2000) and the receptive fields of mammalian V1 complex cells, with either random-bar stimuli aligned to the preferred orientation of the cell (Touryan et al., 2002; Rust et al., 2005) or natural images presented in a random temporal sequence (Fig. 6A) (Touryan et al., 2005). When used with natural images, this method needs to be modified to correct for the spatial correlations in the images (Field, 1987; Dong and Atick, 1995; Simoncelli and Olshausen, 2001).

In addition to STC, complex-cell receptive fields have been analyzed with a phase-separated Fourier model (see below, *Evaluating what we know about V1 and beyond*), supervised training of neural networks (Lehky et al., 1992; Lau et al., 2002; Prenger et al., 2004), least-square algorithms (Ringach et al., 2002), and information maximization (Sharpee et al., 2004). Here we will focus the discussion on the results of the STC analysis.

For most of the complex cells in cat V1, STC identified two significant eigenvectors (Touryan et al., 2002, 2005), each of which corresponds to a stimulus pattern (referred to as a “feature”) that is effective for driving the cell. Each feature exhibits ON and OFF subregions, resembling the receptive fields of simple cells (Hubel and Wiesel, 1962; Jones and Palmer, 1987; DeAngelis et al., 1993a). The spatial profiles of the significant eigenvectors along the axis perpendicular to the preferred orientation of the cell can be well approximated by Gabor functions, and the two significant eigenvectors of each cell exhibit similar spatial frequencies but a phase difference of $\sim 90^\circ$ (Fig. 6B).

After identifying these visual features for each cell, the next step is to quantify the contribution of each feature to the response of the cell. This is achieved by computing the contrast–response function of each significant eigenvector (Chichilnisky, 2001; Touryan et al., 2002). The contrast of the eigenvector in each stimulus is measured as the dot product of the eigenvector and the stimulus, and the contrast–response function for the eigenvector is computed as the average firing rate of the cell at each contrast of the eigenvector. For all complex cells, the firing rate was found to increase with eigenvector contrast at both positive and negative polarities (Fig. 6C), consistent with the known polarity invariance of complex cells (Hubel and Wiesel, 1962).

The bimodal contrast–response functions, together with the spatial phase relationship between the pair of significant eigenvectors for each cell, are consistent with the well known energy

Average firing rate is plotted against the contrast of each eigenvector shown in **B**. Error bar indicates $\pm$ SEM. Dashed lines, Fits of the data with the function $r(x) = \beta x^\gamma + r_0$, where r is the firing rate, x is contrast, and β , γ , and r_0 are free parameters. **D**, **E**, Prediction of cortical responses to natural images. Correlation coefficients between the predicted and measured responses based on the eigenvectors were plotted against those based on the linear receptive field (**D**) and against the estimated upper bound (**E**). Each symbol represents one complex cell.

model for complex cells (Adelson and Bergen, 1985). The energy model combines the outputs of a quadrature pair of subunits to produce an orientation-selective and spatial frequency-selective but phase-invariant complex-cell receptive field (Fig. 1C). Mathematically, the model can be described as $R = F^\phi(k^\phi * s) + F^{\phi + 90^\circ}(k^{\phi + 90^\circ} * s)$, where R is the response of the neuron, s is the stimulus, k^ϕ is the receptive field of a subunit (ϕ represents preferred spatial phase), and F represents the contrast–response function of the subunit, which is commonly approximated by $F(x) = x^2$. The pair of significant eigenvectors identified in the STC analysis appear to correspond nicely to k^ϕ and $k^{\phi + 90^\circ}$. Note that the quadratic nonlinearity of the energy model in fact provides the perfect substrate for the STC analysis, which computes the second-order Wiener kernel of the cell. Thus, the consistency between the STC result and the energy model is not entirely surprising. Interestingly, however, artificial neural networks that are trained to approximate the input–output relationship of complex cells also tend to converge to connection profiles that closely resemble the energy model (Lau et al., 2002; Prenger et al., 2004), suggesting that, given the structural constraint of feedforward neural networks, the energy model is especially suited for approximating the responses of complex cells.

Although the STC analysis on the complex cells of the cat has yielded results consistent with the energy model, the result in monkey V1 is more complicated. In the monkey, STC analysis consistently revealed additional excitatory and even suppressive eigenvectors for some complex cells, suggesting the existence of excitatory and suppressive influences beyond those predicted by the energy model (Rust et al., 2005). The structure of the additional excitatory eigenvectors was consistent with a model in which complex cells are constructed by the convergence of a number of spatially shifted subunits. The suppressive eigenvectors had a primarily divisive influence on the excitatory response and was consistent with a weighted normalization mechanism. The difference between the results in cat and monkey may be attributed to the species difference or to certain technical factors, such as the amount of data available for the analysis.

The studies described above have led to relatively compact descriptions of complex-cell receptive fields. Based on these descriptions, how much can we predict the responses of the neuron to arbitrary stimuli? Here, we will first consider the responses to simple stimuli that are commonly used to measure the response properties of visual neurons. We will then address the issue of complex stimuli including natural scenes.

Direction selectivity

In studies using random-bar stimuli, the subunit receptive fields of many complex cells consist of ON and OFF subregions shifting smoothly over time (spatiotemporally inseparable receptive field), suggesting direction selectivity (Lau et al., 2002; Rust et al., 2005). Previous studies have shown that direction selectivity of simple cells can be predicted from their spatiotemporal receptive fields and the expansive nonlinearity in the contrast–response function (Albrecht and Geisler, 1991; DeAngelis et al., 1993b; Heeger, 1993). Because the responses of complex cells are approximated as the sum of two or more subunits, one can predict their direction selectivity using the receptive field of each subunit and its contrast–response function. Lau et al. (2002) made the prediction based on the subunit receptive fields identified with neural networks and found that the prediction agreed reasonably well with the direction selectivity measured with drifting sinusoidal gratings. The correlation coefficient between the predicted and actual direction selectivity was ~ 0.8 , which is considerably

higher than that based on the linear receptive field estimated by simple spike-triggered averaging (correlation coefficient, $r = 0.47$). Of course, this prediction is not perfect, and it is unclear yet how much the prediction error is attributable to errors in the estimated model parameters and how much is attributable to other types of nonlinearities not captured by the model.

Orientation selectivity

When complex-cell receptive fields are mapped with two-dimensional stimuli, the spatial structure of their subunits clearly suggests orientation selectivity (Fig. 6B). One can again predict the orientation tuning of the cell based on the subunit receptive fields and their contrast–response functions. In a study in cat V1 (Toussyn et al., 2005), the predicted tuning curve agreed well with the tuning curve measured with drifting gratings, in both the preferred orientation (absolute difference, 3.6°) and the tuning bandwidth (mean absolute difference, 6.3°). The dot product between the predicted and the measured tuning curves was 0.95 ± 0.04 (SD). Note that, in this prediction, the nonlinear transformation from intracellular signal to spike output is reflected in the contrast–response function (Fig. 6C), which helps to prevent overestimation of the orientation tuning bandwidth (see above, *Direction selectivity*).

Spatial frequency tuning

The subunit receptive fields can also be used to derive the spatial frequency tuning, which was tested in the STC analysis in cat V1 (Toussyn et al., 2005). However, the result is confounded by the fact that the receptive fields were mapped with natural images, which requires a modification of the STC method that corrects for the spatial correlations in natural images. Because the stimulus power at high spatial frequencies is relatively low for natural images (Field, 1987; Dong and Atick, 1995; Simoncelli and Olshausen, 2001), this correction tends to amplify noise in the high-frequency range (Theunissen et al., 2000). To reduce noise amplification, the correction is made only at spatial frequencies below a cutoff point, which directly affects the predicted spatial frequency tuning of the cell. Thus, how well the complex-cell receptive fields measured with STC can predict the spatial frequency tuning of the cells remains to be examined.

Complex stimuli

As an ultimate validation of the STC model for complex cells, the model is used to predict the responses to complex stimulus ensembles, including random bars and natural images, following the same procedure as that used for predicting direction and orientation tuning. To compare the predicted and measured firing rates, an important limiting factor is the variability in the measured firing rate attributable to the limited number of repeats of the stimulus sequence. The upper bound of the correlation coefficient between the predicted and measured signals attributable to measurement variability can be estimated from the measured responses (Hsu et al., 2004). When the correlation coefficient between the predicted and the measured responses is plotted against this estimated upper bound, it is clear that, although the prediction based on the STC model is consistently better than that based on simple spike-triggered averaging (Fig. 6D), it is far below the upper bound for the majority of complex cells (Fig. 6E).

There may be several factors that affect the performance of the model. First, there is likely to be inaccuracy of the model because of limited amount of data. For example, there may be additional subunits that are not identified by the STC analysis (Rust et al., 2005). The accuracy of the measured eigenvectors and their con-

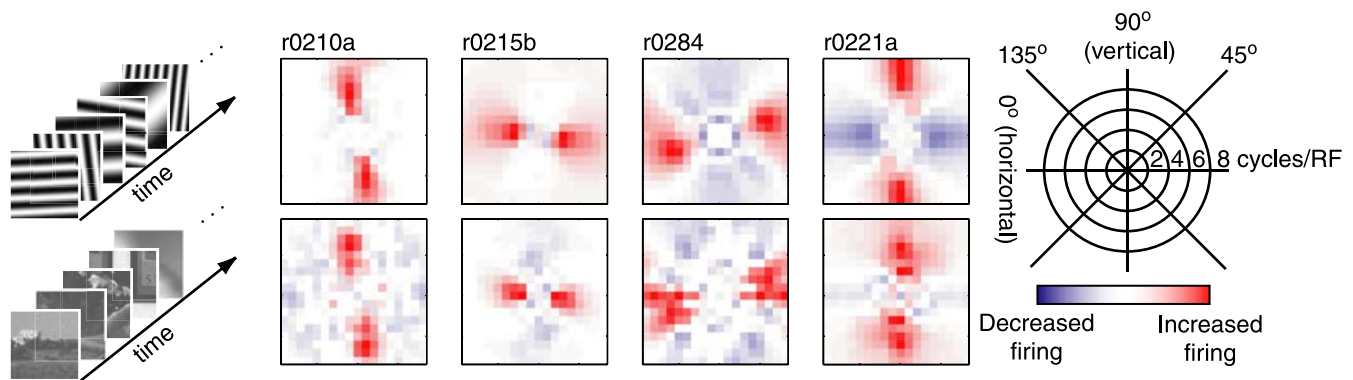


Figure 7. Spatial receptive fields of four V1 neurons estimated using dynamic grating sequences (top row) and dynamic sequences of natural images (bottom row). The spatial receptive field describes selectivity in terms of a joint orientation-spatial frequency tuning surface. Each point in a spatial receptive field map describes the relative response to a stimulus of a particular orientation and spatial frequency (angle about the origin and distance from the origin, respectively; see legend at right). Red indicates orientations and spatial frequencies that tended to increase responses; blue indicates decreased responses. Both gratings and natural images yield spatial receptive fields in which excitatory tuning is centered around a small range of spatial frequencies and orientations. Inhibitory tuning tends to be more diffuse, and its structure depends on the stimulus.

trast–response functions also depends on the amount of data. Second, cortical cells are known to exhibit several other types of nonlinearity, including contrast adaptation (Maffei et al., 1973), gain control (Heeger, 1992a), and contextual modulation by stimuli outside of the classical receptive field (Fitzpatrick, 2000; Freeman et al., 2001). These effects are clearly not captured by the model that sums the responses of two subunits, and they can certainly contribute to the failure of the STC model in predicting the responses to complex stimuli. In the study in cat V1, although the receptive fields measured with natural images were similar to those measured with random stimuli, natural images were found to be more effective for driving the cells, resulting in a higher gain in the contrast–response functions (Felsen et al., 2005). In a study in monkey V1 (David et al., 2004), the spatiotemporal receptive field measured with a phase-separated Fourier model depends significantly on the statistics of visual stimuli (Fig. 7) (see below, *Evaluating what we know about V1 and beyond*). Because these nonlinear effects are not yet described in a compact and precise form, it remains a formidable challenge to build a “universal model” for complex cells that can predict their responses to arbitrary stimulus ensembles.

Evaluating what we know about V1 and beyond

Computational studies of primary visual cortex have produced powerful quantitative models that accurately describe neuronal responses to simple stimuli (Daugman, 1980; Adelson and Bergen, 1985; Carandini et al., 1997). Recent experiments suggest that, under natural viewing conditions, neuronal responses may deviate markedly from the predictions of established models (Smyth et al., 2003; David et al., 2004; Touryan et al., 2005). Are these deviations functionally important? This section describes a nonlinear system identification (NLSI) approach that can address this problem. This approach has already been used to determine precisely how well current models predict neuronal responses during natural vision (Theunissen et al., 2001; David et al., 2004; Prenger et al., 2004; David and Gallant, in press). It can also be used to identify the underlying causes of poor predictions and to determine their relative importance.

Most quantitative neuronal models of visual processing are based on neurophysiological data gathered using simple stimuli such as bars, gratings, and spots of light. Simple stimulus experiments provide a powerful means to test specific hypotheses about neuronal function. However, if neurons function differ-

ently under natural viewing conditions, then models based on simple stimulus experiments might not be accurate. To determine how well such models account for natural visual responses, they must be used to make quantitative predictions about natural visual responses. Then these predictions must be tested in neurophysiological experiments under conditions approximating natural vision. It is therefore essential to have an objective method for comparing visual models.

NLSI provides a general and efficient method for testing the predictions of neuronal models under natural viewing conditions (Theunissen et al., 2001). Within the NLSI framework, each neuron is a nonlinear filter that operates on the sensory input to produce an output (a series of action potentials). The filter is estimated from neurophysiological data and expressed as an equation that describes the functional response properties of the neuron.

NLSI analysis requires several steps: model specification, fitting and regularization, and validation and evaluation. First, a computational model is specified that describes the range of potential filtering operations that a neuron might implement. In principle, one can use any quantitative model that instantiates a specific theory of neuronal function or a nonparametric model estimated directly from the data. In practice, it is common to use a very simple and general model that is efficient to estimate. Second, a regression procedure is used to fit the model to neurophysiological data. Many fitting algorithms are available; the optimal algorithm for a specific problem depends on several factors, including the complexity of the stimulus, the stimulus and response sample sizes, neuronal noise, and the nature of the computational model. In neurophysiological experiments on natural vision, a successful fit will depend critically on the regularization procedure used to reduce the influence of neuronal noise (Theunissen et al., 2001).

The third step of NLSI is validation and evaluation of model predictions. Cortical neurons are quite variable in their responses, and this variability will tend to cause overfitting. Models that have many free parameters will tend to fit both systematic variability and experimental noise. This overfitting problem makes it difficult to compare models with different numbers of free parameters. However, this problem can be eliminated by using a strict cross-validation procedure (Theunissen et al., 2001; David et al., 2004; David and Gallant, in press). Before analysis, a portion of the data (typically 5–10%) is set aside for validation.

The model is then fit to the remaining data. Performance of the model is assessed by evaluating its ability to predict responses in the reserved validation dataset.

The performance of a model is usually expressed as the correlation between predicted and observed responses in the validation dataset (Theunissen et al., 2001; David et al., 2004; David and Gallant, in press), although other measures have also been used (Hsu et al., 2004). One important consideration when evaluating such correlations is the intrinsic variability of the data. Neuronal noise and the data sample size place an upper limit on the correlation between any model and a specific dataset (Hsu et al., 2004; David and Gallant, in press). In any real experiment, even the best possible model cannot produce a perfect correlation between predicted and observed responses. The “potentially explainable variance” is the fraction of total response variance that could theoretically be predicted, given neuronal noise and a finite sample. The predictive power of a model is the percentage of potentially explainable variance that is, in fact, explained by the model.

Although the NLSI approach has been used in neurophysiology for several decades, early studies only used white noise or simple stimuli with flat power spectra (DeBoer and Kuyper, 1968; Sutter, 1975; Emerson et al., 1987; Jones and Palmer, 1987). Because the statistics of natural stimuli are not white (Field, 1987; Dong and Atick, 1995), a nonlinear neuron might respond differently to these simple stimuli than it would to natural images. Fortunately, recent theoretical and experimental studies have shown that NLSI can also be used with natural images (Theunissen et al., 2001; Willmore and Smyth, 2003; David et al., 2004; David and Gallant, in press).

Early NLSI studies usually used simple first- or second-order Volterra/Wiener regression models (Marmarelis and Marmarelis, 1978; Emerson et al., 1987; Eggermont, 1993). These first- and second-order polynomial models are computationally tractable and simple to fit. However, more complicated models that account for nonlinear mechanisms such as contrast gain control (Carandini et al., 1997) and nonclassical receptive-field modulation (Knierim and van Essen, 1992; Vinje and Gallant, 2002) should, in theory, perform better than a simple first- or second-order model.

Several studies have now used some variant of the NLSI approach to investigate how cortical neurons encode natural images (Lehky et al., 1992; Theunissen et al., 2001; Ringach et al., 2002; Smyth et al., 2003; Willmore and Smyth, 2003; David et al., 2004; Prenger et al., 2004; David and Gallant, 2005; Touryan et al., 2005). David et al. (2004) evaluated the performance of current V1 models by examining their ability to predict natural visual responses. The stimuli were movies that simulated the spatial and temporal stimulation that a single neuron would receive during free inspection of a static, monochromatic natural scene. Neurophysiological recordings were made from single V1 neurons during presentation with these natural vision movies. David et al. (2004) developed a spatiotemporal phase-separated Fourier model (PSFT) model that can account for many of the known properties of V1 neurons. According to this model, the spatial response of any V1 neuron is given by the weighted linear sum of stimulus energy across orientation, spatial frequency, and phase channels, followed by an output nonlinearity. The temporal responses are modeled as a simple linear filter. Because the PSFT model incorporates both first- and second-order terms, it can account for both the phase-sensitive responses of V1 simple cells and the phase-independent responses of complex cells (De Valois et al., 1982). The model describes both excitatory and inhibitory interactions between spatial frequency channels, and it can ap-

proximate neuronal mechanisms mediating contrast gain control (Carandini et al., 1997) and nonclassical modulation (Knierim and van Essen, 1992; Vinje and Gallant, 2002). Finally, the PSFT can capture the second-order mechanisms revealed by the Wiener (Emerson et al., 1987) and spike-triggered covariance methods (see above, *Understanding V1 complex cells*) (Rust et al., 2005; Touryan et al., 2005).

David et al. (2004) used a linearized reverse-correlation procedure to fit the PSFT model to the data acquired from each neuron (a software toolbox for conducting this analysis is available at <http://strfpak.berkeley.edu>). They found that spatiotemporal receptive fields estimated using the PSFT model predict ~20% of the total variance in responses of V1 neurons during simulated natural vision. In a later study (David and Gallant, 2005), they measured the intrinsic noise level in these experiments to determine explainable variance and predictive power. They found that a second-order Fourier power (FP) model (see below) predicts ~40% of the explainable variance in the responses of V1 neurons during natural vision. This 40% figure represents the best current estimate of how well conventional V1 models account for natural visual responses.

The study by David et al. (2004) also included two additional stimulus classes: rapid sequences of random gratings that had white spatial and temporal statistics, and rapid sequences of natural images that had the 1/f spatial statistics of natural vision movies and the white temporal statistics of grating sequences. A separate spatiotemporal receptive field was estimated from the responses to each stimulus class, giving three (potentially different) filter estimates for each neuron. They reported that receptive field generated using data from a single stimulus class predict responses within that stimulus class better than across classes. Temporal stimulus statistics have a large effect on estimated receptive field; temporal responses become more bimodal and short-term adaptation increases as the temporal stimulus statistics become more natural (less white). Spatial stimulus statistics have a more subtle effect on the receptive field; excitatory tuning appears to be stable regardless of the spatial stimulus structure, but natural spatial statistics evoke complex changes in the pattern of inhibitory tuning. Together, these observations demonstrate that the functional properties of V1 neurons depend partly on the prevailing stimulus statistics (Fig. 7).

David et al. (2004) reported that the modest performance of the PSFT model was most likely attributable to the unanticipated influence of temporal stimulus statistics on V1 response properties. Temporal stimulus statistics dramatically change the temporal integration properties of these neurons, and a linear temporal filter cannot account for this temporal nonlinearity. Current research (J. L. Gallant, unpublished observations) suggests that a more sophisticated temporally nonlinear model will substantially improve predictions. To improve model performance beyond this will require a still more complicated model that explicitly represents other nonlinear mechanisms such as the spatial interaction between the classical and nonclassical receptive field (Knierim and van Essen, 1992; Vinje and Gallant, 2002).

Most previous NLSI studies have focused on the peripheral nervous system or primary sensory cortex; work on extrastriate visual areas such as V4 is in its infancy. One current study (Gallant, unpublished observations) uses a quantitative FP model that describes orientation tuning, spatial frequency tuning, and position invariance reported previously in V4 (Desimone and Schein, 1987; Gallant et al., 1996). Recordings were made from single V4 neurons during stimulation with a 4 Hz sequence of natural images (Hayden and Gallant, 2005). A linearized reverse-

correlation procedure was used to fit the FP model to the data acquired from each V4 neuron. The estimated spatial receptive fields predict on average ~10% of the explainable variance of spatial responses in V4 (Gallant, unpublished observations). These predictions are substantially worse than those obtained in area V1, although the V4 study did not attempt to model the temporal receptive field. This result confirms that relatively little is currently known about the functional properties of neurons in extrastriate visual areas along the ventral pathway.

The predictive power of any neuronal model will depend on how accurately the model captures the nonlinear stimulus–response mapping function of a neuron. Because there are currently no commonly accepted models of processing in V4, it is not surprising that the FP model performs poorly. One strategy for developing an appropriate model is to use NLSI to generate the model itself. These nonparametric models are not specified explicitly but instead emerge from the data during the fitting procedure. Several nonparametric models have been used to characterize V1 neurons: artificial neural networks (Lehky et al., 1992; Lau et al., 2002; Prenger et al., 2004), maximally informative dimensions (Sharpee et al., 2004), and kernel regression methods such as the support vector machine (Wu and Gallant, 2004). Because these models do not require any previous theory about the nature of neural coding, they are also likely to be useful for characterizing neurons in visual areas beyond V1. One complication of nonparametric models is that the estimated spatio-temporal receptive field may be broadly distributed across many coefficients in such a way that it cannot be interpreted directly. In such cases, a separate visualization procedure can be used to interpret the receptive field (Lau et al., 2002; Prenger et al., 2004).

Responses in extrastriate areas such as V4 are strongly modulated by attention (Connor et al., 1997; Gallant, 2003), but most neurophysiological studies of attentional modulation are conducted under conditions much simpler than those prevailing during natural vision. One recent study (David et al., 2002) investigated how attention modulates neuronal responses in area V4 during a naturalistic free-viewing visual search task (Mazer and Gallant, 2003). The FP model and linearized reverse correlation were used to estimate the spatial receptive field of each V4 neuron. A stepwise regression procedure was then used to identify attentional modulation of the mean rate, response gain, and orientation and spatial frequency tuning. David et al. (2002) found that attention modulates all three aspects of the spatial receptive field; it changes mean response rate and gain, and it modulates orientation and spatial frequency tuning. Additional investigation of more naturalistic experimental conditions will be crucial to obtain an accurate understanding the role of attention during natural vision.

The ultimate test of any theory of the neural basis of visual perception is its ability to predict neuronal responses during natural vision. NLSI provides an efficient framework for developing testable models, fitting them to neurophysiological data and evaluating their predictive power. This process also provides a means to compare models objectively based on their ability to predict responses during natural vision. Such comparisons can facilitate model selection and identify promising areas for future research. The available evidence (David et al., 2004; David and Gallant, 2005) indicates that current models predict ~40% of the explainable variance in responses of V1 neurons during natural vision. Preliminary data suggest that there remains much to learn about the functional characteristics of neurons beyond V1. In extrastriate visual areas such as V4, most of the knowable is still unknown.

What we don't know about V1

The past 40 years have produced enormous amounts of data concerning the structure and function of area V1 in a variety of different species of mammals. What has emerged from this work, among other things, is a fairly well agreed on standard model of V1 neuron response properties, usually involving a combination of linear filtering, half-wave rectification and squaring, and response normalization (described above in previous sections). Although this model is well supported by much of the available data, it is still unknown how well it fares in accounting for the actual behavior of the entire population of V1 neurons when presented with the full complexity of time-varying natural scenes.

In previous work, Olshausen and Field (2005) attempted to quantify our current level of understanding of V1 function by considering two important factors: an estimate of the fraction of V1 neuron types that are typically characterized in experimental studies and the fraction of variance explained in the responses of these neurons under natural viewing conditions. Together, these two factors led them to conclude that, at present, as much as 85% of V1 function has yet to be accounted for. They identified five specific problems that will need to be overcome before we can claim to have an accurate, standard model of V1 function. These are briefly reviewed here.

Biased sampling of neurons

The vast majority of our knowledge about V1 function has been obtained from single-unit recordings in which a microelectrode is brought into close proximity with a neuron in cortex. Ideally, when doing this, one would like to obtain an unbiased sample from any given layer of cortex, but some biases are difficult to avoid. The most troubling of these is that the process of hunting for neurons with a single microelectrode will typically steer one toward neurons with higher firing rates. Recent studies of energy consumption by neurons estimate that the average activity of neurons be relatively low, i.e., <1 spike/s in primate cortex (Attwell and Laughlin, 2001; Lennie, 2003). However, one finds many studies in the literature in which even the spontaneous or background rates are well above 1 spike/s, suggesting that the more active neurons are substantially overrepresented (Lennie, 2003). Olshausen and Field (2005) estimate that as much as 60% of the population of neurons in V1 may have been missed because of this bias. A number of recent studies show that, when one searches for neurons using less biased methods, such as chronic implants or antidromic stimulation, neurons with substantially lower firing rates become much more common (for review, see Olshausen and Field, 2004). If the same is true for V1, then we will need to characterize the response properties of such neurons, and, if they are substantially different, we may need to revisit our beliefs about how “typical” V1 neurons behave.

Biased stimuli

Many elements of the existing standard model for V1 neurons were derived from experiments using a fairly restricted class of test stimuli. Oftentimes, these stimuli are ideal for characterizing linear systems, i.e., spots, white noise, or sine-wave gratings, or else they are designed around preexisting notions of how neurons should respond. The hope is that the insights gained from studying neurons using these reduced stimuli will generalize to more complex situations, i.e., natural scenes. However, in a nonlinear system, the response to any reduced set of stimuli cannot be guaranteed to provide the information needed to predict the response to an arbitrary combination of those stimuli. The extent to which this holds for V1 has yet to be thoroughly examined. Be-

cause it is impossible to map out the response to all possible stimuli, some assumptions about the nature of the nonlinearity and the stimulus space must be made. The assumption that Olshausen and Field (2005) believe is appropriate is that the nonlinearities relevant to visual processing are most likely to be revealed when the system is presented with ecologically relevant stimuli. Traditionally, experimentalists have been reluctant to use natural scenes as stimuli because they seem highly variable and “uncontrolled.” However, in recent years, there has been significant progress in modeling the structure of natural images (Simoncelli and Olshausen, 2001), and it should soon be possible to develop parametric descriptions of natural images that could be used to generate experimental stimuli (Heeger and Bergen, 1995). Several recently developed adaptive stimulus techniques also provide a promising avenue for determining the relevant stimulus for sensory neurons (Edin et al., 2004; Foldiak et al., 2004; O’Connor et al., 2004).

Biased theories

Currently in neuroscience there is an emphasis on “telling a story.” This often encourages investigators to demonstrate when a theory explains data, not when a theory provides a poor model. In addition, editorial pressures can encourage one to make a tidy picture out of data that may actually be quite messy. The result is that theories emerge that are centered around explaining a particular subset of published data or that can be conveniently proven rather than being motivated by functional considerations. A good theory should not only explain data, but it must also address how the problems of vision are solved by the cortex. For example, much of our thinking about V1 function has been guided by the notion that there are two distinct classes of cells, simple and complex, but a number of recent studies are now calling this into question, pointing out that this classification scheme could simply be an artifact of the lens through which we view the data (Mechler and Ringach, 2002; Priebe et al., 2004). Indeed, given the variety of response properties one observes, it can become quite difficult to shoehorn any given cell into one of these two categories (see above, *Understanding V1 simple cells*). Another theory bias often embedded in investigations of V1 function is the notion that simple cells and complex cells are actually coding for the presence of edges, corners, or other two-dimensional shape features in images. However, despite much effort in computer vision, it has proven impossible to detect even the simple outline of an object using a filter such as a simple- or complex-cell model. Moreover, it is entirely unclear whether such a representation would be meaningful or useful in the first place. One of the most challenging problems facing the cortex is that of inferring a representation of three-dimensional surfaces from the two-dimensional image (Nakayama et al., 1995). This is not an easy problem to solve, and it still lies beyond the capabilities of modern computer vision. It seems quite likely that V1 plays a role in solving this problem, but understanding how it does so might require going beyond bottom-up filtering models to consider how top-down information is used in the interpretation of images (Lee and Mumford, 2003).

Interdependence and contextual effects

A combination of anatomical and physiological studies suggest that ~60–80% of the response variance of a layer 4 V1 neuron is a function of other V1 neurons or inputs other than those arising from LGN (for review, see Olshausen and Field, 2005). Determining how these contextual signals influence the response prop-

erties of V1 neurons has been the subject of many investigations over the past decade, often using bars or gratings to probe how stimuli in the surround affect the response to a stimulus in the center of the receptive field (for review, see Albright and Stoner, 2002; Series et al., 2003). However, the problem one faces in teasing apart contextual effects this way is the combinatorial explosion in exploring all of the possible spatial and featural configurations of surrounding stimuli. What we really want to know is how neurons respond within the sorts of context encountered in natural scenes. Some of the initial studies exploring the role of context in natural scenes have demonstrated pronounced nonlinear effects that tend to sparsify activity in a way that would have been hard to predict from the existing studies (Vinje and Gallant, 2000). More studies along these lines are needed, and, most importantly, we need to understand how and why the context in natural scenes produces such effects.

Ecological deviance

In the past few years, a number of laboratories have begun using natural scenes as stimuli when recording from neurons in the visual pathway. In particular, Gallant and collaborators have taken the approach of attempting to determine how well one can predict the responses of V1 neurons to natural stimuli using a variety of different models. For example, David et al. (2004) have explored two different types of models: a linearized spatiotemporal receptive field model, in which the response of the neuron is essentially a weighted sum of the image pixels over space and time, and a phase-separated Fourier model that allows one to capture the phase-invariance nonlinearity of a complex cell (see above, *Evaluating what we know about V1*). These models can typically explain between 30 and 40% of the response variance of V1 neurons. One could possibly obtain a better fit to the data by including additional terms modeling suppression (Rust et al., 2005) and temporal adaptation (Lesica et al., 2003), or even a spiking mechanism (Paninski et al., 2004), but it is still sobering to realize that the receptive field component per se, which is the bread and butter of the standard model, accounts for so little of the response variance. Moreover, the way in which these models fail does not leave one optimistic that the addition of modulatory terms or pointwise nonlinearities will fix matters. Typically, the model will undershoot the response of the neuron, but there are also many events in the response that are completely missed by the model and vice versa. This is in stark contrast to the LGN, in which the linear model predicts nearly every event in the response but mainly differs by a gain factor that can often be corrected by the proper application of a gain control mechanism (see above, *Understanding LGN responses*) (Fig. 3). Also, in an experiment in ferrets using brief presentations of static natural images, the linear model often appears to succeed in conveying the gist of neural responses (Smyth et al., 2003) (see above, *Understanding V1 simple cells*). Thus, there appears to be a qualitative mismatch in predicting the responses of cortical neurons to time-varying natural images that will require more than tweaking to resolve. What seems to be suggested by the data is that a more complex, network nonlinearity is at work here and that describing the behavior of any one neuron will require one to include the influence of other simultaneously recorded neurons.

Obtaining a more complete and accurate picture of V1 function will require three fundamental changes in our approach: (1) more widespread use of recording techniques that allow for an unbiased sample of the population of neurons across different layers of cortex and their interactions, (2) the use of time-varying, natural scenes, or reasonable approximations thereof, for char-

acterizing neural response properties, and (3) the advancement of functionally driven, and testable, theories of V1 function. The latter will require that we extrapolate beyond the available data, taking into account the problems that need to be solved by the visual system in addition to constraints provided by known neuroanatomy and neurophysiology. Importantly, we will need to keep an open mind in considering new theories, and, given how little of V1 function we can currently claim to understand, we should be prepared for some surprises as new data come in.

Conclusions

Among the views expressed in this review, there are a number of points of agreement, but there are also notable differences of opinion.

A first point of agreement is that an adequate model of visual responses should predict responses to arbitrary stimuli, not only those encountered in the laboratory but also those seen in nature. Surprisingly, many of the standard models of early visual processing have not been held to this rigorous test. As pointed out by Demb (see above, *Understanding the retinal output*), models of the retina have not attempted to go all the way and predict responses to complex video sequences. This attempt has been made in LGN, with a linear model by Dan et al. (1996) and with a nonlinear model described in *Understanding LGN responses*. As described in the subsequent sections, some initial attempts have been made for V1 cells. Still, much work remains before we can tell whether current models are appropriate to predict responses to complex video sequences.

A second point of agreement is that a tractable model of visual responses should include a linear receptive field (or more than one, in the case of V1 complex cells). This linear receptive field could operate directly on a scaled version of the images (as in the models derived from those in Fig. 1) or on a transformation of the stimulus such as the Fourier transform, as is the case for the model described by Gallant (see above, *Evaluating what we know about V1 and beyond*). In fact, even models of V1 complex cells that postulate highly nonlinear image processing eventually combine the visual information with a linear receptive field (Ringach et al., 2002). Similarly, an influential model of visual responses in the cortical middle temporal area postulates a linear receptive field whose inputs are the outputs of V1 neurons (Simoncelli and Heeger, 1998).

A point at which the approaches differ is in the use of complex stimuli such as natural video sequences. One approach, proposed above in *Understanding LGN responses*, is to use them only to test a model that has been constrained with simpler stimuli. An alternative approach, proposed above in *Evaluating what we know about V1 and beyond* and *What we don't know about V1*, is to use them also to discover the appropriate type of model and to constrain the model. There are valid reasons for both approaches.

The first approach posits that appropriate models of neural function are so nonlinear that it would be hopeless to try to fit them to responses to complicated stimuli. It is better to constrain the model with simpler stimuli such as gratings and then see how the model does with more complicated stimuli. This approach has illustrious precedent in biophysics: for example, Hodgkin and Huxley characterized the mechanisms of the action potential by using highly simplified stimuli (e.g., by clamping the voltage of the cell). They did not try to inject a natural-looking current that simulates the arrival of synaptic potentials; even with the computers of today, this approach would be unlikely to yield Hodgkin and Huxley's elegant equations.

However, neurons in visual cortex, and particularly in areas

beyond V1, are likely to be specialized in scene analysis that goes well beyond the extraction of edges and similar low-level image processing. It could become hopeless to try to characterize these neurons using simple stimuli such as gratings. Stimuli of that simplicity might perhaps become useful for this task later, after a wide exploration is made with complex, natural stimuli and the general outlines of the mechanisms underlying the responses have been elucidated.

Another point at which there does not seem to be agreement is whether we have a satisfactory grasp of the computations performed in primary visual cortex. Although no study yet seems to have measured how well a model of retinal ganglion cell would do in predicting responses to a natural movie, at least for the more linear X-type ganglion cells, there is a sense that a model that included the key known nonlinear mechanisms should do fairly well (see above, *Understanding the retinal output*). Similarly, the model of LGN presented above (see above, *Understanding LGN responses*), although far from perfect, does capture the gist of LGN responses to complex stimuli. Models of neurons in visual cortex succeed in capturing the qualitative (tuning) properties of these neurons but are less robust at quantitatively predicting the responses to arbitrary stimulus sets. *Understanding V1 simple cells* describes a fairly successful attempt made for V1 simple cells, but that attempt only involved static images. The model of complex cells described in *Understanding V1 complex cells* can be derived directly from responses to complex natural video sequences (Touryan et al., 2005), but the model does not perform as well as models of earlier visual neurons. The measurements presented in *Evaluating what we know about V1 and beyond* indicate that a model that combines linear and second-order mechanisms accounts for at least one-third of the variance of V1 responses to complex stimuli. However, we do not know how well a more complete model would do in predicting V1 responses: it does not seem that anybody has yet attempted to build a model of V1 neurons that includes many of the known dynamical nonlinear components (e.g., synaptic depression and surround suppression) (Fig. 5) and tested this model on responses to complex video sequences. Until this work is done, the field is open to the criticisms levied above in *What we don't know about V1*.

Three obstacles lie in the way of a meaningful comparison of different models for a given visual stage or across visual stages. First, as long as different models are applied to responses of different stimuli (flashed still pictures vs cartoons vs natural video sequences), we will not be able to compare model performance. Second, there does not seem to be agreement over the timescale of responses that one is trying to describe. Some models attempt to predict responses down to the individual spike (Keat et al., 2001; Paninski et al., 2004), but more commonly one is interested in firing rates computed in the 10 ms range (see above, *Understanding V1 simple cells*). Third, there does not seem to be an agreed measure of model quality. For example, the different sections of this review invoked different measures, including correlation (Tolhurst), percentage of the variance (Demb and Mante), and percentage of the explainable variance or explainable correlation (Gallant and Dan). The last approaches seem the most reasonable, because they include an estimate of the variability of responses, which the models are not expected to capture. Recent work indeed has yielded revised measures of percentage of variance (Sahani and Linden, 2003) or of correlation (Hsu et al., 2004) that are adjusted to indicate when a model accounts for the explainable responses, i.e., when the deviations from the actual responses are within the variability that is present in the responses.

Clearly, much work lies ahead before we can say that we understand what the early visual system does. Recent advances in stimulus presentation and receptive-field mapping techniques allow us, for the first time, to fit and test models that can produce quantitative predictions of the response of an individual neuron to large classes of stimuli. Although these models appear to perform reasonably (albeit imperfectly) in the retina and LGN, the performance of these models degrades substantially as one ascends to the cortex. Standard models of neurons in the retina (Fig. 1A) successfully describe both the qualitative characteristics of the tuning properties of the neuron (e.g., tuning for the spatial frequency of a drifting grating) and are reasonable predictors of the response of a neuron at brief timescales (accounting for ~80% of the variance in response; see above, *Understanding the retinal output*). Similarly, standard models of V1 simple cells (Fig. 1B) successfully capture the basic tuning properties (orientation, direction, and phase sensitivity) of the “most linear” neurons falling in this class. In the more nonlinear V1 simple and complex cells, response properties can no longer be described by a standard model including a single linear filter and only recently have new characterization techniques allowed researchers to recover models of these cells. Models of V1 complex cells (Fig. 1C) quite successfully predict the qualitative response properties of these neurons, such as tuning for the orientation and direction of moving stimuli. However, standard models of these neurons succeed in capturing only ~35% of the explainable variance in natural visual responses (see above, *Evaluating what we know about V1 and beyond*), suggesting that crucial elements are missing from the standard model description.

One likely source of error in standard models is the absence of any history-dependent adaptation. At all stages of processing, visual neurons adapt to the recent luminance and contrast of stimuli; these properties cannot be captured by the models presented in Figure 1. As demonstrated in *Understanding LGN responses*, standard models can be extended to include these adaptive properties, and this extension is important for predicting the responses to complex, naturalistic stimuli. Inclusion of these dynamic mechanisms in models of cortical neurons, in which adaptation effects are known to be even more severe than at earlier stages, will likewise improve their predictive power.

Extending these characterization techniques to include other known dynamic nonlinearities will also improve their performance. For example, including a realistic (non-Poisson) spike generator that captures a refractory period and integrate-and-fire dynamics significantly improves the predictability of the response at short timescales (e.g., <50 ms) (Keat et al., 2001; Bair and Movshon, 2004; Paninski et al., 2004).

Among all these limitations, a clear point of agreement that emerges from this review is the need for functional models that provide a compact description of the transformation of stimuli into the response of the neuron. Functional models are not necessarily (and perhaps preferably not) directly mapped onto the known biophysics and anatomy. They are composed of idealized boxes such as linear filters, divisive stages, and nonlinearities, all simple components that can provide a compact answer to the question “What does this neuron compute?” This question is distinct from the question “How does this neuron give this response?” but the two questions are clearly related. Just as knowing the Hodgkin-Huxley equations has greatly helped the discovery of how ion channels work, knowing which computations a neuron performs on visual images can act as a powerful guide to understanding the underlying biology. In addition to guiding the investigation of underlying biological mechanism, a successful

functional model for a visual stage is required if we want to understand computation at later stages, and indeed, a functional model is what is needed to establish the link between neural activity and perception, which is a central goal of sensory neuroscience.

References

- Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2:284–299.
- Albrecht DG, Geisler WS (1991) Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis Neurosci* 7:531–546.
- Albrecht DG, Hamilton DB (1982) Striate cortex of monkey and cat: contrast response function. *J Neurophysiol* 48:217–237.
- Albright TD, Stoner GR (2002) Contextual influences on visual processing. *Annu Rev Neurosci* 25:339–379.
- Andrews BW, Pollen DA (1979) Relationship between spatial frequency selectivity and receptive field profile of simple cells. *J Physiol (Lond)* 287:163–176.
- Atick JJ, Redlich AN (1990) Towards a theory of early visual processing. *Neural Comput* 2:308–320.
- Attwell D, Laughlin SB (2001) An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab* 21:1133–1145.
- Baccus SA, Meister M (2002) Fast and slow contrast adaptation in retinal circuitry. *Neuron* 36:909–919.
- Bair W, Movshon JA (2004) Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *J Neurosci* 24:7305–7323.
- Baylor DA, Hodgkin AL, Lamb TD (1974) Reconstruction of the electrical responses of turtle cones to flashes and steps of light. *J Physiol (Lond)* 242:759–791.
- Benardete EA, Kaplan E (1999) The dynamics of primate M retinal ganglion cells. *Vis Neurosci* 16:355–368.
- Berry MJ, Warland DK, Meister M (1997) The structure and precision of retinal spike trains. *Proc Natl Acad Sci USA* 94:5411–5416.
- Berson DM (2003) Strange vision: ganglion cells as circadian photoreceptors. *Trends Neurosci* 26:314–320.
- Betsch BY, Einhauser W, Kording KP, König P (2004) The world from a cat's perspective—statistics of natural videos. *Biol Cybern* 90:41–50.
- Blakemore C, Tobin EA (1972) Lateral inhibition between orientation detectors in the cat's visual cortex. *Exp Brain Res* 15:439–440.
- Bonds AB (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis Neurosci* 2:41–55.
- Bonin V, Mante V, Carandini M (2005) The suppressive field of neurons in lateral geniculate nucleus. *J Neurosci*, in press.
- Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. *Neuron* 26:695–702.
- Brodie SE, Knight BW, Ratliff F (1978) The spatiotemporal transfer function of the Limulus lateral eye. *J Gen Physiol* 72:167–203.
- Cai D, DeAngelis GC, Freeman RD (1997) Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. *J Neurophysiol* 78:1045–1061.
- Carandini M (2004) Receptive fields and suppressive fields in the early visual system. In: *The cognitive neurosciences III*, Ed 3 (Gazzaniga MS, ed), pp 313–326. Cambridge, MA: MIT.
- Carandini M, Ferster D (2000) Membrane potential and firing rate in cat primary visual cortex. *J Neurosci* 20:470–484.
- Carandini M, Heeger DJ (1994) Summation and division by neurons in primate visual cortex. *Science* 264:1333–1336.
- Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci* 17:8621–8644.
- Carandini M, Heeger DJ, Movshon JA (1999) Linearity and gain control in V1 simple cells. In: *Cerebral cortex*, Vol 13, Models of cortical circuits (Ulinski PS, Jones EG, Peters A, eds), pp 401–443. New York: Kluwer Academic/Plenum.
- Carandini M, Heeger DJ, Senn W (2002) A synaptic explanation of suppression in visual cortex. *J Neurosci* 22:10053–10065.
- Chander D, Chichilnisky EJ (2001) Adaptation to temporal contrast in primate and salamander retina. *J Neurosci* 21:9904–9916.
- Cheng H, Chino YM, Smith III EL, Hamamoto J, Yoshida K (1995) Transfer

- characteristics of lateral geniculate nucleus X neurons in the cat: effects of spatial frequency and contrast. *J Neurophysiol* 74:2548–2557.
- Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. *Network* 12:199–213.
- Chisum HJ, Fitzpatrick D (2004) The contribution of vertical and horizontal connections to the receptive field center and surround in V1. *Neural Netw* 17:681–693.
- Cleland BG, Enroth-Cugell C (1968) Quantitative aspects of sensitivity and summation in the cat retina. *J Physiol (Lond)* 198:17–38.
- Cleland BG, Freeman AW (1988) Visual adaptation is highly localized in the cat's retina. *J Physiol (Lond)* 404:591–611.
- Connor CE, Preddie DC, Gallant JL, Van Essen DC (1997) Spatial attention effects in macaque area V4. *J Neurosci* 17:3201–3214.
- Dacey DM, Liao HW, Peterson BB, Robinson FR, Smith VC, Pokorny J, Yau KW, Gamlin PD (2005) Melanopsin-expressing ganglion cells in primate retina signal colour and irradiance and project to the LGN. *Nature* 433:749–754.
- Dan Y, Atick JJ, Reid RC (1996) Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *J Neurosci* 16:3351–3362.
- Daugman JG (1980) Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Res* 20:847–856.
- David SV, Gallant JL (2005) Predicting neuronal responses during natural vision. *Network*, in press.
- David SV, Mazer JA, Gallant JL (2002) Pattern-specific attentional modulation of V4 spatiotemporal receptive fields during free-viewing visual search. *Soc Neurosci Abstr* 31:559.2.
- David SV, Vinje WE, Gallant JL (2004) Natural stimulus statistics alter the receptive field structure of v1 neurons. *J Neurosci* 24:6991–7006.
- Dawis S, Shapley R, Kaplan E, Tranchina D (1984) The receptive field organization of X-cells in the cat: spatiotemporal coupling and asymmetry. *Vision Res* 24:549–564.
- De Ruyter Van Steveninck R, Bialek W (1988) Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. *Proc R Soc Lond Ser B Biol Sci* 234:379–414.
- De Valois RL, De Valois K (1988) *Spatial vision*. Oxford: Oxford UP.
- De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res* 22:545–559.
- DeAngelis GC, Robson JG, Ohzawa I, Freeman RD (1992) Organization of suppression in receptive fields of neurons in cat visual cortex. *J Neurophysiol* 68:144–163.
- DeAngelis GC, Ohzawa I, Freeman RD (1993a) Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J Neurophysiol* 69:1091–1117.
- DeAngelis GC, Ohzawa I, Freeman RD (1993b) Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *J Neurophysiol* 69:1118–1135.
- Dean AF, Tolhurst DJ (1983) On the distinctness of simple and complex cells in the visual cortex of the cat. *J Physiol (Lond)* 344:305–325.
- Dean AF, Tolhurst DJ (1986) Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat striate cortex. *Exp Brain Res* 62:143–151.
- DeBoer E, Kuyper P (1968) Triggered correlation. *IEEE Trans Biomed Eng* 15:159–179.
- Demb JB (2002) Multiple mechanisms for contrast adaptation in the retina. *Neuron* 36:781–783.
- Demb JB, Haarsma L, Freed MA, Sterling P (1999) Functional circuitry of the retinal ganglion cell's nonlinear receptive field. *J Neurosci* 19:9756–9767.
- Demb JB, Zaghloul K, Haarsma L, Sterling P (2001) Bipolar cells contribute to nonlinear spatial summation in the brisk-transient (Y) ganglion cell in mammalian retina. *J Neurosci* 21:7447–7454.
- Demb JB, Sterling P, Freed MA (2004) How retinal ganglion cells prevent synaptic noise from reaching the spike output. *J Neurophysiol* 92:2510–2519.
- Derrington AM, Lennie P (1984) Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *J Physiol (Lond)* 357:219–240.
- Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57:835–868.
- Dong DW, Atick JJ (1995) Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus. *Network* 6:159–178.
- Edin F, Machens CK, Schutze H, Herz AV (2004) Searching for optimal sensory signals: iterative stimulus reconstruction in closed-loop experiments. *J Comput Neurosci* 17:47–56.
- Eggermont JJ (1993) Wiener and Volterra analyses applied to the auditory system. *Hear Res* 66:177–201.
- Emerson RC, Citron MC, Vaughn WJ, Klein SA (1987) Nonlinear directionally selective subunits in complex cells of cat striate cortex. *J Neurophysiol* 58:33–65.
- Enroth-Cugell C, Freeman AW (1987) The receptive-field spatial structure of cat retinal Y cells. *J Physiol (Lond)* 384:49–79.
- Enroth-Cugell C, Jakiela HG (1980) Suppression of cat retinal ganglion cell responses by moving patterns. *J Physiol (Lond)* 302:49–72.
- Enroth-Cugell C, Robson JG (1966) The contrast sensitivity of retinal ganglion cells of the cat. *J Physiol (Lond)* 187:517–522.
- Enroth-Cugell C, Shapley RM (1973a) Adaptation and dynamics of cat retinal ganglion cells. *J Physiol (Lond)* 233:271–309.
- Enroth-Cugell C, Shapley RM (1973b) Flux, not retinal illumination, is what cat retinal ganglion cells really care about. *J Physiol (Lond)* 233:311–326.
- Enroth-Cugell C, Lennie P, Shapley RM (1975) Surround contribution to light adaptation in cat retinal ganglion cells. *J Physiol (Lond)* 247:579–588.
- Felsen G, Touryan J, Han F, Dan Y (2005) Cortical sensitivity to visual features in natural scenes. *PLoS Biol* 3:e342.
- Ferster D (1988) Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. *J Neurosci* 8:1172–1180.
- Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A* 4:2379–2394.
- Field DJ, Tolhurst DJ (1986) The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proc R Soc Lond B Biol Sci* 228:379–400.
- Fitzpatrick D (2000) Seeing beyond the receptive field in primary visual cortex. *Curr Opin Neurobiol* 10:438–443.
- Foldiak P, Xiao D, Keyser C, Edwards R, Perrett DI (2004) Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Prog Brain Res* 144:107–116.
- Freeman RD, Ohzawa I, Walker G (2001) Beyond the classical receptive field in the visual cortex. *Prog Brain Res* 134:157–170.
- Freeman T, Durand S, Kiper D, Carandini M (2002) Suppression without inhibition in visual cortex. *Neuron* 35:759–771.
- Fuortes MGF, Hodgkin AL (1964) Changes in the time scale and sensitivity in the ommatidia of *Limulus*. *J Physiol (Lond)* 172:239–263.
- Gallant JL (2003) Neural mechanisms of natural scene perception. In: *The visual neurosciences* (Chalupa LM, Werner JS, eds). Boston: MIT.
- Gallant JL, Connor CE, Rakshit S, Lewis JW, Van Essen DC (1996) Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76:2718–2739.
- Gardner JL, Anzai A, Ohzawa I, Freeman RD (1999) Linear and nonlinear contributions to orientation tuning of simple cells in the cat's striate cortex. *Vis Neurosci* 16:1115–1121.
- Glezer VD, Tsherbach TA, Gauseman VE, Bondarko VM (1982) Spatiotemporal organization of receptive fields of the cat striate cortex. The receptive fields as the grating filters. *Biol Cybern* 43:35–49.
- Graham NVS (1989) *Visual pattern analyzers*. New York: Oxford UP.
- Guido W, Weyand T (1995) Burst responses in thalamic relay cells of the awake behaving cat. *J Neurophysiol* 74:1782–1786.
- Guillery RW, Sherman SM (2002) Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. *Neuron* 33:163–175.
- Hayden BY, Gallant JL (2005) Timecourse of attention reveals different mechanisms for spatial and feature attention in V4. *Neuron* 47:637–643.
- Heeger DJ (1992a) Normalization of cell responses in cat striate cortex. *Vis Neurosci* 9:181–197.
- Heeger DJ (1992b) Half-squaring in responses of cat striate cells. *Vis Neurosci* 9:427–443.
- Heeger DJ (1993) Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J Neurophysiol* 70:1885–1898.
- Heeger DJ, Bergen JR (1995) Pyramid based texture analysis/synthesis. In: *Proceedings of the 22nd annual conference on computer graphics and interactive techniques*, pp 229–238. New York: ACM.

- Hirsch JA (2003) Synaptic physiology and receptive field structure in the early visual pathway of the cat. *Cereb Cortex* 13:63–69.
- Hirsch JA, Alonso JM, Reid RC, Martinez LM (1998) Synaptic integration in striate cortical simple cells. *J Neurosci* 18:9517–9528.
- Hochstein S, Shapley RM (1976) Linear and nonlinear spatial subunits in Y cat retinal ganglion cells. *J Physiol (Lond)* 262:265–284.
- Hosoya T, Baccus SA, Meister M (2005) Dynamic predictive coding by the retina. *Nature* 436:71–77.
- Hsu A, Borst A, Theunissen FE (2004) Quantifying variability in neural responses and its application for the validation of model predictions. *Network* 15:91–109.
- Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J Physiol (Lond)* 148:574–591.
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol (Lond)* 160:106–154.
- Jagadeesh B, Wheat HS, Ferster D (1993) Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science* 262:1901–1904.
- Jones HE, Andolina IM, Oakley NM, Murphy PC, Sillito AM (2000) Spatial summation in lateral geniculate nucleus and visual cortex. *Exp Brain Res* 135:279–284.
- Jones JP, Palmer LA (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol* 58:1187–1211.
- Kaplan E, Marcus S, So YT (1979) Effects of dark adaptation on spatial and temporal properties of receptive fields in cat lateral geniculate nucleus. *J Physiol (Lond)* 294:561–580.
- Kaplan E, Purpura K, Shapley R (1987) Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus. *J Physiol (Lond)* 391:267–288.
- Kara P, Reinagel P, Reid RC (2000) Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron* 27:635–646.
- Kayser C, Salazar receptive field, Konig P (2003) Responses to natural scenes in cat V1. *J Neurophysiol* 90:1910–1920.
- Keat J, Reinagel P, Reid RC, Meister M (2001) Predicting every spike: a model for the responses of visual neurons. *Neuron* 30:803–817.
- Kim KJ, Rieke F (2001) Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. *J Neurosci* 21:287–299.
- Kim KJ, Rieke F (2003) Slow Na^+ inactivation and variance adaptation in salamander retinal ganglion cells. *J Neurosci* 23:1506–1516.
- Knierim JJ, van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J Neurophysiol* 67:961–980.
- Kremers J, Weiss S, Zrenner E (1997) Temporal properties of marmoset lateral geniculate cells. *Vision Res* 37:2649–2660.
- Kuffler SW (1953) Discharge patterns and functional organization of mammalian retina. *J Neurophysiol* 16:37–68.
- Kuffler SW, Fitzhugh R, Barlow HB (1957) Maintained activity in the cat's retina in light and darkness. *J Gen Physiol* 40:683–702.
- Lampl I, Anderson JS, Gillespie DC, Ferster D (2001) Prediction of orientation selectivity from receptive field architecture in simple cells of cat visual cortex. *Neuron* 30:263–274.
- Lankheet MJ, Van Wezel RJ, Prickaerts JH, van de Grind WA (1993a) The dynamics of light adaptation in cat horizontal cell responses. *Vision Res* 33:1153–1171.
- Lankheet MJ, Przybyszewski AW, van de Grind WA (1993b) The lateral spread of light adaptation in cat horizontal cell responses. *Vision Res* 33:1173–1184.
- Lau B, Stanley GB, Dan Y (2002) Computational subunits of visual cortical neurons revealed by artificial neural networks. *Proc Natl Acad Sci USA* 99:8974–8979.
- Lee BB, Dacey DM, Smith VC, Pokorny J (2003) Dynamics of sensitivity regulation in primate outer retina: the horizontal cell network. *J Vis* 3:513–526.
- Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am A Opt Image Sci Vis* 20:1434–1448.
- Lehky SR, Sejnowski TJ, Desimone R (1992) Predicting responses of nonlinear neurons in monkey striate cortex to complex patterns. *J Neurosci* 12:3568–3581.
- Lennie P (2003) The cost of cortical computation. *Curr Biol* 13:493–497.
- Lesica NA, Stanley GB (2004) Encoding of natural scene movies by tonic and burst spikes in the lateral geniculate nucleus. *J Neurosci* 24:10731–10740.
- Lesica NA, Bolorio AS, Stanley GB (2003) Adaptive encoding in the visual pathway. *Network* 14:119–135.
- Li B, Peterson MR, Thompson JK, Duong T, Freeman RD (2005) Cross-orientation suppression: monoptic and dichoptic mechanisms are different. *J Neurophysiol* 94:1645–1650.
- Maffei L, Fiorentini A (1976) The unresponsive regions of visual cortical receptive fields. *Vision Res* 16:1131–1139.
- Maffei L, Fiorentini A, Bisti S (1973) Neural correlate of perceptual adaptation to gratings. *Science* 182:1036–1038.
- Mante V (2005) Gain controls based on luminance and contrast in the early visual system. In: Institute of neuroinformatics. PhD thesis. Federal Institute of Technology, Zurich.
- Mante V, Bonin V, Carandini M (2004) How local contrast determines gain and temporal integration in LGN. *Soc Neurosci Abstr* 30:409.10.
- Mante V, Frazor RA, Bonin V, Geisler WS, Carandini M (2005) Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci*, in press.
- Marmarelis PZ, Marmarelis VZ (1978) Analysis of physiological systems: the white noise approach. New York: Plenum.
- Martinez LM, Wang Q, Reid RC, Pillai C, Alonso JM, Sommer FT, Hirsch JA (2005) Receptive field structure varies with layer in the primary visual cortex. *Nat Neurosci* 8:372–379.
- Masland RH (2001) Neuronal diversity in the retina. *Curr Opin Neurobiol* 11:431–436.
- Mata ML, Ringach DL (2005) Spatial overlap of ON and OFF subregions and its relation to response modulation ratio in macaque primary visual cortex. *J Neurophysiol* 93:919–928.
- Mazer JA, Gallant JL (2003) Goal-related activity in V4 during free viewing visual search. Evidence for a ventral stream visual salience map. *Neuron* 40:1241–1250.
- McLean J, Palmer LA (1989) Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Res* 29:675–679.
- Mechler F, Ringach DL (2002) On the classification of simple and complex cells. *Vision Res* 42:1017–1033.
- Movshon JA, Thompson ID, Tolhurst DJ (1978a) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol (Lond)* 283:53–77.
- Movshon JA, Thompson ID, Tolhurst DJ (1978b) Receptive field organization of complex cells in the cat's striate cortex. *J Physiol (Lond)* 283:79–99.
- Mukherjee P, Kaplan E (1995) Dynamics of neurons in the cat lateral geniculate nucleus: in vivo electrophysiology and computational modeling. *J Neurophysiol* 74:1222–1243.
- Nakayama K, He ZJ, Shimojo S (1995) Visual surface representation: a critical link between lower-level and higher-level vision. In: An invitation to cognitive science (Kosslyn SM, Osherson DN, eds), pp 1–70. Cambridge, MA: MIT.
- O'Connor KN, Petkov CI, Sutter ML (2004) Stimulus optimization for auditory cortical neurons. *Soc Neurosci Abstr* 30:529.514.
- Olshausen BA, Field DJ (2004) Sparse coding of sensory inputs. *Curr Opin Neurobiol* 14:481–487.
- Olshausen BA, Field DJ (2005) How close are we to understanding v1? *Neural Comput* 17:1665–1699.
- Olveczky BP, Baccus SA, Meister M (2003) Segregation of object and background motion in the retina. *Nature* 423:401–408.
- Ozeki H, Sadakane O, Akasaki T, Naito T, Shimegi S, Sato H (2004) Relationship between excitation and inhibition underlying size tuning and contextual response modulation in the cat primary visual cortex. *J Neurosci* 24:1428–1438.
- Palmer LA, Davis TL (1981) Receptive-field structure in cat striate cortex. *J Neurophysiol* 46:260–276.
- Paninski L, Pillow JW, Simoncelli EP (2004) Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model. *Neural Comput* 16:2533–2561.
- Passaglia CL, Enroth-Cugell C, Troy JB (2001) Effects of remote stimulation on the mean firing rate of cat retinal ganglion cells. *J Neurosci* 21:5794–5803.
- Pillow JW, Simoncelli EP (2003) Biases in white noise analysis due to non-Poisson spike generation. *Neurocomputing* 52:109–115.

- Prenger R, Wu MC, David SV, Gallant JL (2004) Nonlinear V1 responses to natural scenes revealed by neural network analysis. *Neural Netw* 17:663–679.
- Priebe NJ, Mechler F, Carandini M, Ferster D (2004) The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nat Neurosci* 7:1113–1122.
- Purpura K, Tranchina D, Kaplan E, Shapley RM (1990) Light adaptation in the primate retina: analysis of changes in gain and dynamics of monkey retinal ganglion cells. *Vis Neurosci* 4:75–93.
- Ramcharan EJ, Gnadt JW, Sherman SM (2005) Higher-order thalamic relays burst more than first-order relays. *Proc Natl Acad Sci USA* 102:12236–12241.
- Reich DS, Victor JD, Knight BW, Ozaki T, Kaplan E (1997) Response variability and timing precision of neuronal spike trains in vivo. *J Neurophysiol* 77:2836–2841.
- Reid RC, Soodak RE, Shapley RM (1991) Directional selectivity and spatio-temporal structure of receptive fields of simple cells in cat striate cortex. *J Neurophysiol* 66:505–529.
- Reid RC, Victor JD, Shapley RM (1997) The use of m-sequences in the analysis of visual neurons: linear receptive field properties. *Vis Neurosci* 14:1015–1027.
- Rieke F (2001) Temporal contrast adaptation in salamander bipolar cells. *J Neurosci* 21:9445–9454.
- Ringach DL, Hawken MJ, Shapley R (2002) Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *J Vis* 2:12–24.
- Robson JG (1975) Receptive fields: spatial and intensive representation of the visual image. In: *Handbook of perception* (Carterette D, Friedman W, eds). New York: Academic.
- Rodiek RW (1998) *The first steps in seeing*. Sunderland, MA: Sinauer.
- Rodiek RW, Brenning RK, Watanabe M (1993) The origin of parallel visual pathways. In: *Contrast sensitivity* (Shapley RM, Man-Kit Lam D, eds). Cambridge, MA: Bradford Books.
- Roska B, Werblin F (2001) Vertical interactions across ten parallel, stacked representations in the mammalian retina. *Nature* 410:583–587.
- Rust NC, Schwartz O, Movshon JA, Simoncelli EP (2005) Spatiotemporal elements of macaque V1 receptive fields. *Neuron* 46:945–956.
- Sahani M, Linden JF (2003) How linear are auditory cortical responses? In: *Advances in neural information processing systems* (Becker S, Thrun S, Obermeyer K, eds), pp 301–308. Cambridge, MA: MIT.
- Saito H, Fukada Y (1986) Gain control mechanisms in X- and Y-type retinal ganglion cells of the cat. *Vision Res* 26:391–408.
- Sakai HM, Naka K (1995) Response dynamics and receptive-field organization of catfish ganglion cells. *J Gen Physiol* 105:795–814.
- Saul AB, Humphrey AL (1990) Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus. *J Neurophysiol* 64:206–224.
- Schumer RA, Movshon JA (1984) Length summation in simple cells of cat striate cortex. *Vision Res* 24:565–571.
- Sclar G (1987) Expression of “retinal” contrast gain control by neurons of the cat’s lateral geniculate nucleus. *Exp Brain Res* 66:589–596.
- Sclar G, Maunsell JHR, Lennie P (1990) Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res* 30:1–10.
- Sengpiel F, Vorobyov V (2005) Intracortical origins of interocular suppression in the visual cortex. *J Neurosci* 25:6394–6400.
- Sengpiel F, Baddeley RJ, Freeman TC, Harrad R, Blakemore C (1998) Different mechanisms underlie three inhibitory phenomena in cat area 17. *Vision Res* 38:2067–2080.
- Series P, Lorenceau J, Fregnac Y (2003) The “silent” surround of V1 receptive fields: theory and experiments. *J Physiol (Paris)* 97:453–474.
- Shapley R, Lennie P (1985) Spatial frequency analysis in the visual system. *Annu Rev Neurosci* 8:547–583.
- Shapley R, Victor JD (1979) Nonlinear spatial summation and the contrast gain control of cat retinal ganglion cells. *J Physiol (Lond)* 290:141–161.
- Shapley RM, Enroth-Cugell C (1984) Visual adaptation and retinal gain controls. In: *Progress in retinal research* (Osborne N, Chader G, eds.), pp 263–346. London: Pergamon.
- Shapley RM, Victor JD (1978) The effect of contrast on the transfer properties of cat retinal ganglion cells. *J Physiol (Lond)* 285:275–298.
- Shapley RM, Victor JD (1981) How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. *J Physiol (Lond)* 318:161–179.
- Sharpee T, Rust NC, Bialek W (2004) Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Comput* 16:223–250.
- Sherman SM (2001) Tonic and burst firing: dual modes of thalamocortical relay. *Trends Neurosci* 24:122–126.
- Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. *Vision Res* 38:743–761.
- Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- Simoncelli EP, Pillow JW, Paninski L, Schwartz O (2004) Characterization of neural responses with stochastic stimuli. In: *The cognitive neurosciences* (Gazzaniga M, ed), pp 327–338. Cambridge, MA: MIT.
- Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M (1997) Adaptation of retinal processing to image contrast and spatial scale. *Nature* 386:69–73.
- Smith GD, Cox CL, Sherman SM, Rinzal J (2000) Fourier analysis of sinusoidally driven thalamocortical relay neurons and a minimal integrate-and-fire-or-burst model. *J Neurophysiol* 83:588–610.
- Smyth D, Willmore B, Baker GE, Thompson ID, Tolhurst DJ (2003) The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *J Neurosci* 23:4746–4759.
- So YT, Shapley R (1981) Spatial tuning of cells in and around lateral geniculate nucleus of the cat: X and Y relay cells and perigeniculate interneurons. *J Neurophysiol* 45:107–120.
- Solomon SG, White AJ, Martin PR (2002) Extraclassical receptive field properties of parvocellular, magnocellular, and koniocellular cells in the primate lateral geniculate nucleus. *J Neurosci* 22:338–349.
- Sterling P, Demb JB (2004) *Retina*. In: *Synaptic organization of the brain*, Ed 5 (Sheppard G, ed), pp 217–269. New York: Oxford UP.
- Sutter EE (1975) A revised conception of visual receptive fields based on pseudorandom spatiotemporal pattern stimuli. In: *Proceedings of the first symposium on testing and identification of nonlinear systems* (Mc Cann GD, Marmarelis PZ, eds), pp 353–365. Pasadena, CA: California Institute of Technology.
- Tadmor Y, Tolhurst DJ (1989) The effect of threshold on the relationship between the receptive-field profile and the spatial-frequency tuning curve in simple cells of the cat’s striate cortex. *Vis Neurosci* 3:445–454.
- Taylor WR, Vaney DI (2003) New directions in retinal research. *Trends Neurosci* 26:379–385.
- Theunissen FE, Sen K, Doupe AJ (2000) Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J Neurosci* 20:2315–2331.
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL (2001) Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network* 12:289–316.
- Tolhurst DJ, Dean AF (1987) Spatial summation by simple cells in the striate cortex of the cat. *Exp Brain Res* 66:607–620.
- Tolhurst DJ, Dean AF (1990) The effects of contrast on the linearity of spatial summation of simple cells in the cat’s striate cortex. *Exp Brain Res* 79:582–588.
- Tolhurst DJ, Dean AF (1991) Evaluation of a linear model of directional selectivity in simple cells of the cat’s striate cortex. *Vis Neurosci* 6:421–428.
- Tolhurst DJ, Heeger DJ (1997a) Contrast normalization and a linear model for the directional selectivity of simple cells in cat striate cortex. *Vis Neurosci* 14:19–25.
- Tolhurst DJ, Heeger DJ (1997b) Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. *Vis Neurosci* 14:293–309.
- Touryan J, Lau B, Dan Y (2002) Isolation of relevant visual features from random stimuli for cortical complex cells. *J Neurosci* 22:10811–10818.
- Touryan J, Felsen G, Dan Y (2005) Spatial structure of complex cell receptive fields measured with natural images. *Neuron* 45:781–791.
- Troy JB, Enroth-Cugell C (1993) X and Y ganglion cells inform the cat’s brain about contrast in the retinal image. *Exp Brain Res* 93:383–390.
- Troy JB, Robson JG (1992) Steady discharges of X and Y retinal ganglion cells of cat under photopic illuminance. *Vis Neurosci* 9:535–553.
- Uzzell VJ, Chichilnisky EJ (2004) Precision of spike trains in primate retinal ganglion cells. *J Neurophysiol* 92:780–789.
- van Hateren JH, Rüttiger L, Sun H, Lee BB (2002) Processing of natural temporal stimuli by macaque retinal ganglion cells. *J Neurosci* 22:9945–9960.

- Victor J (1987) The dynamics of the cat retinal X cell centre. *J Physiol (Lond)* 386:219–246.
- Victor JD (1979) Nonlinear systems analysis: comparison of white noise and sum of sinusoids in a biological system. *Proc Natl Acad Sci USA* 76:996–998.
- Victor JD, Shapley RM (1979) Receptive field mechanisms of cat X and Y retinal ganglion cells. *J Gen Physiol* 74:275–298.
- Vinje WE, Gallant JL (2000) Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287:1273–1276.
- Vinje WE, Gallant JL (2002) Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *J Neurosci* 22:2904–2915.
- Wässle H (2004) Parallel processing in the mammalian retina. *Nat Rev Neurosci* 5:747–757.
- Watson AB (1987) Efficiency of a model human image code. *J Opt Soc Am A* 4:2401–2417.
- Watson AB, Solomon JA (1997) Model of visual contrast gain control and pattern masking. *J Opt Soc Am A Opt Image Sci Vis* 14:2379–2391.
- Weliky M, Fiser J, Hunt RH, Wagner DN (2003) Coding of natural scenes in primary visual cortex. *Neuron* 37:703–718.
- Wielaard DJ, Shelley M, McLaughlin D, Shapley R (2001) How simple cells are made in a nonlinear network model of the visual cortex. *J Neurosci* 21:5203–5211.
- Willmore B, Smyth D (2003) Methods for first-order kernel estimation: simple-cell receptive fields from responses to natural scenes. *Network* 14:553–577.
- Willmore B, Tolhurst DJ (2001) Characterizing the sparseness of neural codes. *Network* 12:255–270.
- Wu M-C, Gallant JL (2004) What's beyond the second order? Kernel regression techniques for nonlinear functional characterization of visual neurons. *Soc Neurosci Abstr* 30:984.20.
- Yeh T, Lee BB, Kremers J (1996) The time course of adaptation in macaque retinal ganglion cells. *Vision Res* 36:913–931.
- Zaghloul KA, Boahen K, Demb JB (2003) Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *J Neurosci* 23:2645–2654.
- Zaghloul KA, Boahen K, Demb JB (2005) Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. *J Neurosci* 25:860–868.