Topical Review

Mapping receptive fields in primary visual cortex

Dario L. Ringach

Departments of Psychology and Neurobiology, Jules Stein Eye Institute and Brain Research Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA

Nearly 40 years ago, in the pages of this journal, Hubel and Wiesel provided the first description of receptive fields in the primary visual cortex of higher mammals. They defined two classes of cortical cells, 'simple' and 'complex', based on neural responses to simple visual stimuli. The notion of a hierarchy of receptive fields, where increasingly intricate receptive fields are constructed from more elementary ones, was introduced. Since those early days we have witnessed the birth of quantitative methods to map receptive fields and mathematical descriptions of simple and complex cell function. Insights gained from these models, along with new theoretical concepts, are refining our understanding of receptive field structure and the underlying cortical circuitry. Here, I provide a brief historical account of the evolution of receptive field mapping in visual cortex along with the associated conceptual advancements, and speculate on the shape novel theories of the cortex may take as a result these measurements.

(Received 1 April 2004; accepted after revision 17 May 2004; first published online 21 May 2004)
Correspondence D. L. Ringach: Departments of Psychology, Neurobiology, Jules Stein Eye Institute and Brain Research Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA. Email: dario@ucla.edu

Early work on visual electrophysiology characterized the responses of retinal neurones to the onset of bright and dark spots of light on various locations of the receptive field (Kuffler, 1953; Barlow, 1953; Hartline *et al.* 1956; Hartline & Ratliff, 1957, 1958). These investigators described the spatial structure of receptive fields as being divided into 'on' and 'off' regions. 'On' and 'off' regions of the receptive field were defined depending whether the neurone would respond to the onset of a bright or dark spot, respectively. The temporal discharge patterns were also classified as 'transient' or 'sustained', depending on whether the cell kept firing for long periods during the presence of the stimulus, or if it only produced brief responses at times shortly after the stimulus was switched on (Cleland *et al.* 1971). This body of work provided the classical description of centre–surround organization of retinal receptive fields in the retina, where a central area of one sign ('on' or 'off') is surrounded by a concentric region of the opposite sign.

Hubel and Wiesel pursued a similar line of research when they first started recording from cat visual cortex. In one of their early experiments, they noticed that a cell would discharge in response to a moving line shadow cast by the edge of a slide as it was inserted into the ophthalmoscope (Hubel & Wiesel, 1998). They soon realized that the cell would fire only when the line was oriented within a narrow range. Further measurements confirmed that many other cells were also selective to the orientation of boundaries. Thus, the discovery of orientation tuning – perhaps the major transformation in the organization of receptive fields in the early visual pathway – was accidental. Since that discovery much research has focused on whether there are principled ways of determining those 'features' of the environment to which sensory neurones are selective.

By flashing orientated lines at various locations along the receptive field, Hubel and Wiesel classified cortical neurones into two distinct groups: simple and complex. Simple cells were defined as those whose receptive fields could be divided into separate 'on' and 'off' subregions (Fig. 1*A*), while complex cells were defined by exclusion. They also introduced the concept of a hierarchical organization of receptive fields. According to this proposal, simple cells are constructed first from the convergence of geniculate receptive fields aligned in space to produce the observed elongation of 'on' and 'off' subfields (Fig. 1*B*), and complex cells are subsequently constructed from the convergence of simple-cell receptive fields with similar

 DOI: 10.1113/jphysiol.2004.065771

orientation tuning but varying positions (or spatial phases) to generate a field of 'on–off' responses (Fig. 1*C*). As we discuss below, this initial description of two neuronal classes in primary visual cortex and their associated hierarchical circuitry have shaped both experimental and theoretical studies of cortical function.

Admittedly, much has been achieved by manually mapping receptive fields with simple dots, lines and edges. Using this method, Hubel and Wiesel not only discovered orientation tuning in single neurones, but also described the columnar organization of ocular dominance and orientation preferences in the cortex, and explored the consequences of cortical development during sensory deprivation, among many notable findings (Hubel & Wiesel, 1962, 1968, 1977). However, as evident from the discussion that follows, quantitative methods of receptive field mapping have contributed significantly to refine these initial concepts. The goal of this review is to provide a description of the main methods that have been developed over the years to map receptive fields along with the new theoretical concepts that accompany them.

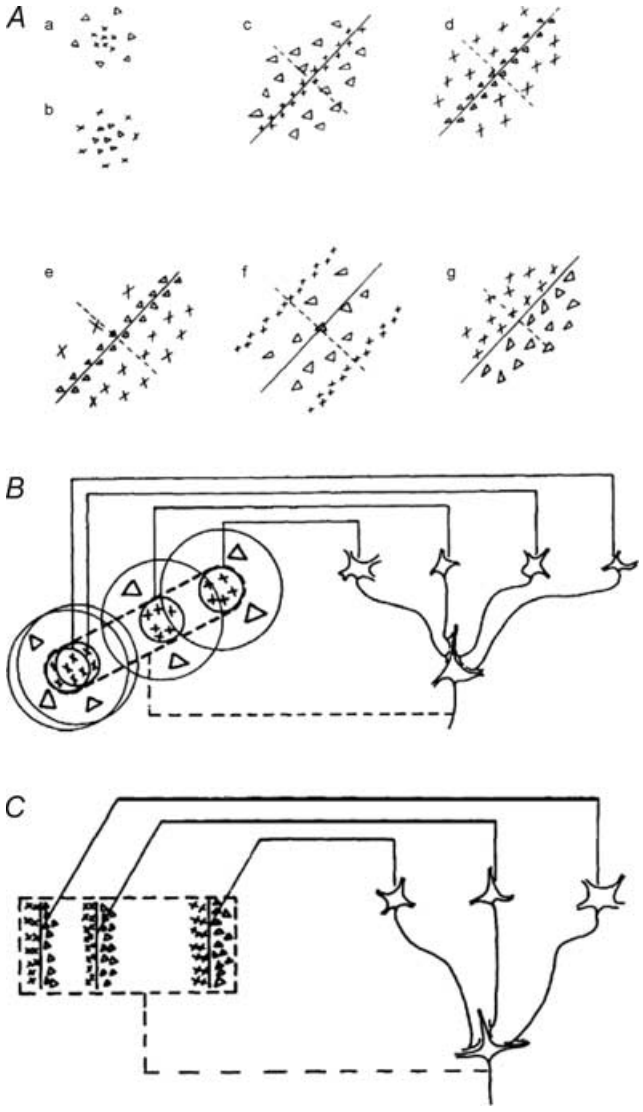

Figure 1. The classical hierarchical model of simple and complex cells

A, simple cells have segregated 'on' and 'off' subregions (examples *A* and *B* are geniculate (receptive fields). *B* and *C*, classical model of simple and complex receptive fields. *B* elongated 'on–off' subregions in simple cells were proposed to be constructed by the convergence of geniculate receptive fields aligned in space. *C*, complex cells are in turn constructed by the convergence of simple cells.

The linear–non-linear model and the reverse correlation method

Arguably, one of the first influential concepts was the idea that some visual receptive fields could be viewed as a spatio-temporal filter acting on the stimulus followed by rectification (the so-called linear–non-linear model), which originated in the study of retinal ganglion cells (Rodieck & Stone, 1965*a*,*b*; Enroth-Cugell & Robson, 1966). Rodieck and Stone successfully used the model to explain the responses of retinal ganglion cells to a variety of static and moving stimuli from the measurements of neural responses to flashing dots. This work yielded an understanding of ganglion cell receptive fields as a linear combination of a centre and surround mechanism that was delayed in time. Similar ideas were, in fact, implicit in the earlier work of Kuffler, Hartline, Barlow and coworkers (Barlow, 1953; Kuffler, 1953; Hartline *et al.* 1956; Hartline & Ratliff, 1958). For example, Kuffler discussed the origin of transient responses as a combination of 'on–off' delayed mechanisms, so he was well aware of the importance of spatio-temporal inseparability of the receptive field (Kuffler, 1953).

Mathematically, the model can be described as follows. The visual stimulus, defined as the spatio-temporal distribution of luminance across the receptive field, is summarized in a vector $\mathbf{x}(t)$ representing the stimulus within the time interval $(t-T, t)$. The assumption is that the neurone has finite memory, so that the response at time t is not influenced by stimuli at times $t' < t-T$. Normally, in primary visual cortex, a selection of $T \approx 300$ ms ensures that all the relevant history of the stimulus is encoded in the vector $\mathbf{x}(t)$. The linear–non-linear model postulates that the spike train is an inhomogeneous Poisson process with instantaneous mean rate given by $\lambda(t) = \varphi(\mathbf{w}^T\mathbf{x})$ (Hunter & Korenberg, 1986; Chichilnisky, 2001; Nykamp & Ringach, 2002; Marmarelis & Marmarelis 1978). Here, $\mathbf{w}^T\mathbf{x}$ describes the linear operation of the filter, and $\varphi(\bullet)$ is a static non-linearity (typically monotonically increasing).

One efficient way to measure the 'kernel', **w**, is by cross-correlation with spatio-temporal white noise (see Chichilnisky, 2001 for a recent review). To the best of my knowledge, the first recording of the full spatio-temporal kernel of a simple cell in cat visual cortex was reported by Erich Sutter in 1975 using a rather elegant apparatus (Fig. 2*A* and *B*) (Sutter, 1975). Because this work appears to be relatively unknown, it is appropriate to briefly describe

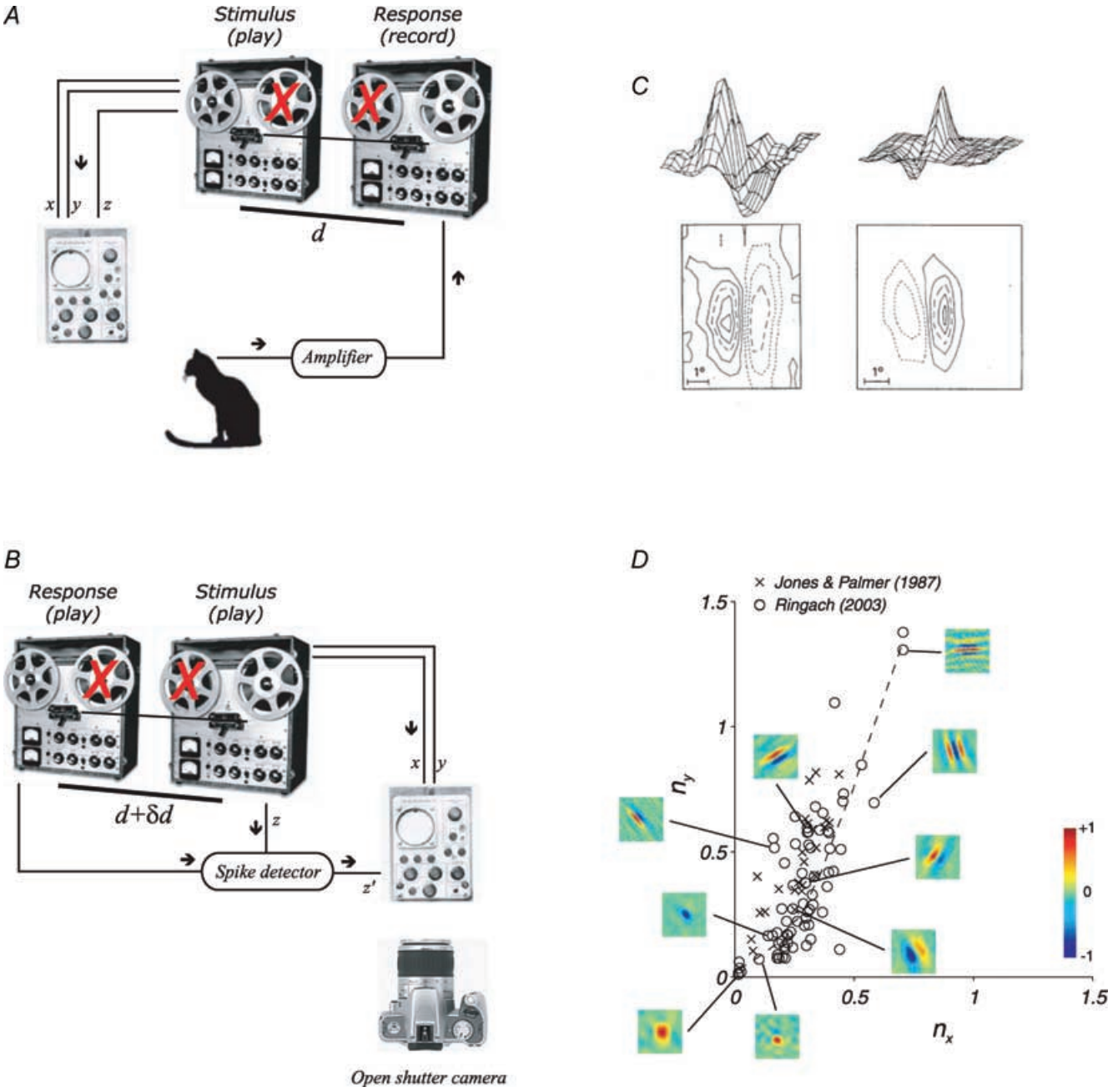

Figure 2. Reverse correlation measurements of simple-cell receptive fields in cat and monkey V1
A and *B*, schematic diagram of the apparatus used by Erich Sutter to measure, for the first time, the full spatio-temporal receptive field of simple cells in cat area 17 (see text for details). *C*, two examples of simple-cell receptive fields measurements from Jones & Palmer (1987*b*). *D*, analysis of the distribution of receptive field shapes in macaque and cat primary visual cortex. The parameter n_x represents the width of the receptive field relative to the period of the underlying grating in a Gabor fit. This number is proportional to the effective number of subregions in the receptive field. Similarly, n_y represents the length (elongation) of the receptive field relative to the period of the underlying grating. Open circles are data from macaque V1, crosses are from cat area 17 (Figure 2*C* reproduced with permission from the American Physiology Society; Jones & Palmer, 1987*a*).

it here. The stimulus was produced by feeding white noise to the position, (x,y), and intensity, z, inputs of an oscilloscope. These signals were generated by playing the different tracks of a single tape (Fig. 2*A*, left recorder). Sutter arranged a second tape recorder (Fig. 2*A*, right recorder) so its head was a distance d apart from the first tape recorder. This second recorder received the tape from the first one and recorded the signals arriving from a microelectrode on another track of the tape. Thus, the temporal relationship between the stimulus and the response was preserved. At any one location on the tape, one can find recorded the neural response and the stimulus that preceded it by d/v seconds. Here, d is the distance between the heads of the tape recorders and v is the linear velocity of the tape.

To analyse the data, the arrangement in Fig. 2*B* was used. The recorder on the left played the audio track with the response of the neurone, while the one of the right played the stimulus. Because the role of the tape recorders is now reversed, and if the distance between the two recorders is d, the audio signals from both tape recorders represent the stimulus and response that occurred at the *same* instant in time. To vary the time lag between the stimulus and the response to be analysed one can simply set the distance at $d + \delta d$, which would represent a time lag of $\tau = \delta d/v$. Once a time lag is selected, the stimulus is re-played on the oscilloscope such that the location signals (x,y) were unchanged, but the intensity signal was multiplied by a brief pulse sequence indicating the presence or absence of a spike in the response audio track. The stimulus was therefore multiplied by the response, and the resulting images were added photographically by having a camera with an open shutter pointed at the oscilloscope screen. One picture was taken for each t and thus the entire spatio-temporal receptive field was computed. This initial study showed the feasibility of applying such methods in visual receptive fields. A careful analysis of their shapes in a large population of cells came at a later stage.

Simple receptive fields have Gabor-like shapes, but not all Gabor-like shapes are receptive fields

Movshon *et al.* (1978) first demonstrated that simple-cortical cells exhibit spatio-temporal summation within their receptive fields. They observed that the spatial profile of the receptive field compared well with synthesized profiles resulting from measurements in response to drifting sinusoidal gratings, indicating that simple-cells showed linear summation up to the spike rectification stage. Jones and Palmer used sparse stimulation with bright and dark 'dots' to measure receptive fields of simple cells in cat area 17 (Jones & Palmer, 1987*b*). By cross-correlating the evoked spikes and with the positions and times of occurrence of the stimuli, they estimated the spatial impulse response at a fixed time lag (50 ms). Some of their original measurements are shown in Fig. 2*C*. This study provided the first measurements of simple-cell receptive field measurements over a population of neurones and established that their shapes were well approximated by two-dimensional Gabor functions (Jones & Palmer, 1987*a*). Using a similar method, the full spatio-temporal kernel in cat area 17 was measured by DeAngelis *et al.*, who also studied their development in kittens (DeAngelis *et al.* 1993*b*).

In macaque primary visual cortex, I used a modification of the reverse correlation method where the receptive field is probed with a fast sequence of gratings and found that the receptive field shapes were similar to those seen in cat (Ringach, 2002). Figure 2*D* shows that the distribution of these shapes in monkey (open circles) and cat (crosses; data from Jones & Palmer, 1987*a*). The parameters n_x and n_y provide a measure proportional to the width and length of the receptive field, respectively, in units of the period of the grating in a two-dimensional Gabor fit to the kernel. Blob-like receptive fields are mapped to points near the origin. Receptive fields with a number of elongated subfields are mapped to points away from the origin. The distribution of (n_x, n_y) appears to lie, approximately, on a one-dimensional curve and cat and monkey data are comparable. The results indicate that a particular family of filter shapes is present in primary visual cortex. Furthermore, current theories of simple-cell receptive fields (Olshausen & Field, 1996; Bell & Sejnowski, 1997; van Hateren & Ruderman, 1998), based on 'optimal' linear representations of the image, fail to account for this distribution (Ringach, 2002). Finding out what is special about this family from a computational point of view may yield clues about their function.

The importance of the output non-linearity and its measurement

Reverse correlation provides a measure of the front-end filter (Chichilnisky, 2001). However, it should be clear that the tuning properties of a linear-nonlinear (LN) system (such as its orientation or spatial frequency bandwidth) do not depend solely on its linear kernel; they will also be influenced by the static non-linearity in the spike generation mechanism, represented by $\varphi(\bullet)$. For example, it has been suggested that both direction and orientation tuning can be sharpened significantly by thresholding or by an accelerating non-linearity (Reid *et al.* 1987, 1991;

Jagadeesh *et al.* 1993; DeAngelis *et al.* 1993*a,b*; Anzai *et al.* 1999). Thus, estimating $\varphi(\bullet)$ is clearly important when comparing the tuning properties of the model to tuning curves obtained with drifting sinusoidal gratings.

One strategy to measure the output non-linearity in the LN system is to perform a two-step analysis (Anzai *et al.* 1999; Chichilnisky, 2001; see also Hunter & Korenberg, 1986). First, measure the linear kernel using reverse correlation, then generate a scatter plot of the linear prediction $\mathbf{w}^T\mathbf{x}$ against the actual response and smooth it to obtain an estimate of the static non-linearity. If we know that the non-linearity can be approximated by some functional family, such as a half power-law rectifier, then a more efficient method can be used where the parameters of the non-linearity are found by matching the input–output moments of the model to the data (Nykamp & Ringach, 2002).

Spatio-temporal inseparable receptive fields and direction selectivity

The concept of the receptive field as a spatio-temporal entity was instrumental in advancing our knowledge of receptive field function (Adelson & Bergen, 1985; Reid *et al.* 1987; Emerson *et al.* 1992; DeAngelis *et al.* 1995). First, it clarified that 'on' and 'off' responses at various locations of the receptive field could be interpreted within the framework of the LN model (DeAngelis *et al.* 1995). Second, it demonstrated how non-trivial computations, such as direction selectivity, may arise from linear mechanisms if the receptive field is inseparable in space and time (Reid *et al.* 1987, 1991).

A receptive field is spatio-temporal separable if it its structure can be written as a product of a spatial and temporal function, $w(x,y,t) = h(x,y)g(t)$. Spatio-temporal inseparability of receptive fields was first shown to be involved in the generation of direction selectivity by measuring the response of the cells with contrast reversing gratings as a function of spatial phase (Enroth-Cugell & Robson, 1966; Hochstein & Shapley, 1976*a,b*; Reid *et al.* 1987). Full measurements of the spatio-temporal receptive fields in both cat and monkey confirmed these early reports and showed a characteristic tilt of 'on' and 'off' subregions in the space–time plane (DeAngelis *et al.* 1993*a,b*). Such tilt in the space–time plane endows the neurone with an asymmetric receptive field that causes the response to be larger when the stimulus moves in one direction than the other. Experimentally, the linear prediction of direction selectivity matches very well the preferred direction of the neurone but underestimates its magnitude by about 1/3 (Reid *et al.* 1991). Intracellular measurements (Jagadeesh *et al.* 1993) have demonstrated that linearity holds very well when one considers the membrane voltage of the neurones, but that thresholding contributes significantly to making the spike responses better tuned than the intracellular voltage demonstrating, once again, the importance of the output non-linearity in the LN system when predicting the tuning properties of spike responses (but see Baker, 2001).

The gain control model

One of the recent advances in the field has been the realization that the LN model of simple cells fails to account for a number of phenomena (Albrecht & Geisler, 1991; Geisler & Albrecht, 1991; Robson, 1991; Heeger, 1992; Carandini *et al.* 1997; Tolhurst & Heeger, 1997). First, measurements of response as a function of orientation show saturation at different levels. This would not be expected from a system where the maximum spike rate saturates due to a fixed output non-linearity. Second, the responses of neurones are suppressed by stimuli that, by themselves, do not cause the cell to fire. A well-documented example is the suppression caused by adding a stimulus orientated orthogonally to the preferred orientation of the cell, a phenomenon referred to as cross-orientation inhibition (Morrone *et al.* 1982, 1987; Bonds, 1989). Third, in response to drifting gratings, the response to increasing contrast does not simply scale but advances in time. In other words, the responses become faster, while a LN model predicts no change in the temporal structure of the response. Fourth, the spatial summation of cells changes with the contrast of the stimulus (Polat *et al.* 1998; Kapadia *et al.* 1999; Sceniak *et al.* 1999); the higher the contrast, the smaller the degree of spatial summation. A LN model would predict no change in spatial summation with contrast. Albrecht *et al.* (2003) provide a recent review of these and other non-linear response properties of cortical neurones.

A number of investigators proposed to extend the LN model by adding a gain control mechanism (Albrecht & Geisler, 1991; Bonds, 1991; Heeger, 1992; Carandini & Heeger, 1994; Tolhurst & Heeger, 1997; Carandini *et al.* 1997). The idea is that the output of the linear filter is divided (or 'normalized') by the overall activity in a population of cortical cells that represent the 'normalization pool' (Fig. 3*A*). The model attributes the selectivity for orientation entirely to the linear filter; it is only the gain that is determined by the normalization signal. The gain control model explains saturation because the activity of the local population, and therefore the normalization signal, increases with

contrast. It also explains cross-orientation inhibition because the normalization signal includes signals that are also tuned to the orthogonal orientation. It has also been recently extended to include the surround from the classical receptive field to explain changes in spatial summation as a function of stimulus contrast (Cavanaugh *et al.* 2002). However, it is likely that a single gain control mechanism may be insufficient to explain the change of contrast–response curves in a variety of suppression phenomena (Sengpiel *et al.* 1998; Carandini *et al.* 2002).

Mathematically, a general form of the gain-control model of a simple cortical cell is:

$$\lambda(t) = \varphi\left(\frac{\mathbf{w}^T\mathbf{x}}{\sqrt{\sigma^2 + \mathbf{x}^T H\mathbf{x}}}\right)$$

Here, $\mathbf{w}^T\mathbf{x}$ describes the linear operation of the filter, we require $\mathbf{x}^T H\mathbf{x} \geq 0$ for all $\mathbf{x}$ (in which case the matrix H is called a positive semidefinite matrix), σ is the semisaturation constant, and $\varphi(\bullet)$ is a static non-linearity usually selected to represent a power-law rectifier: $\varphi(x) = x^\beta$ if $x > 0$ and zero otherwise. A full identification of this model requires that we estimate the linear filter $\mathbf{w}$, the semisaturation constant σ, the exponent β, and the matrix H. The gain control model is attractive because it explains a set of interesting phenomena in a parsimonious way (Carandini *et al.* 1997; Cavanaugh *et al.* 2002). Furthermore, Simoncelli and colleagues have put forward an interesting theoretical framework for gain control (Simoncelli & Olshausen, 2001; Schwartz & Simoncelli, 2001). Within this framework gain control works to increase the degree of independence between neural responses when the system is stimulated with natural signals. It is also worth noting that mechanisms for gain control were also described in classes of retinal (Shapley *et al.* 1972; Shapley & Victor, 1978, 1979) and geniculate neurones (Kaplan *et al.* 1987; Purpura *et al.* 1988; Benardete *et al.* 1992; Benardete & Kaplan, 1999).

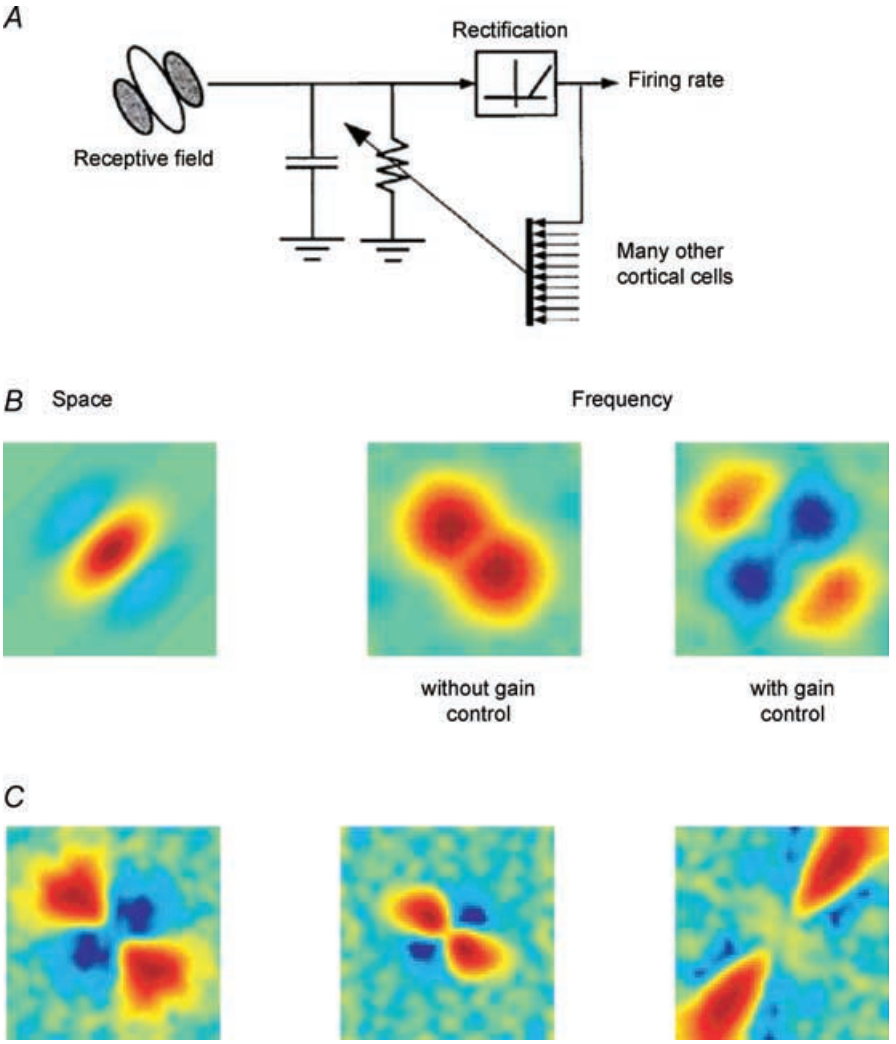

Figure 3. Gain control and sharpening of tuning
A, the gain control model of simple cells. The gain of the front-end filter is divided by the summed activity of a 'normalization pool' (Carandini *et al.* 1997; copyright 1997 by the Society for Neuroscience). *B*, gain control can sharpen tuning in the Fourier domain. The panel on the left shows a simulated Gabor-like receptive field in space. The two panels on the right illustrate the tuning of the neurone in the Fourier (frequency) domain (with the origin at the centre of the panel). The result of gain control is to 'carve away' the activity near the origin resulting in a more localized region that produces response enhancement. The gain control signal is untuned for orientation and low-pass in spatial frequency, nevertheless, it can sharpen tuning. *C*, tuning in the Fourier domain for three sample cells in macaque V1 (Ringach *et al.* 2002). The two examples on the left are consistent with a gain control signal untuned for orientation, but the one on the right shows maximal (net) suppression at oblique angles, implying a tuned suppressive signal (see also Shapley *et al.* 2003 for a review).

One should consider that at least part of the cortical effects may originate from non-linearities in the LGN inputs (see discussion in Carandini *et al.* 1997).

Gain control and intracortical sharpening of tuning

The gain control signal appears to be broadly tuned for orientation, spatial frequency and temporal frequency (DeAngelis *et al.* 1992). It has often been assumed that, because of its broad tuning, the gain control signal cannot sharpen the tuning conferred by the linear filter. This is not entirely correct, and Fig. 3*B* provides an example of how a gain control signal that is untuned in orientation and low-pass in spatial frequency can enhance the tuning of the neurone in the Fourier domain (Ringach *et al.* 2002). The leftmost panel illustrates the spatial kernel of a Gabor receptive field, the two panels on the right show the tuning of the neurone in the Fourier domain in two conditions: with and without gain control. Clearly, the tuning of both spatial frequency and orientation can be enhanced because the gain control signal is 'carving away' the responses near the origin in the Fourier plane (Ringach *et al.* 2002). The result is that the most responsive region in the Fourier domain is shifted away from the origin relative to the response of the linear filter alone. The angular extent of the response enhancement region is reduced. Thus, in cases where the original filter has significant power at low spatial frequencies, gain control can have the net effect of enhancing both spatial frequency and orientation selectivity.

Examples of tuning in the Fourier domain for three macaque V1 cells are shown in Fig. 3*C*. Over the population, such measurements show a correlation between the degree of suppression and tuning in both orientation and spatial frequency, suggesting the same circuitry involved in gain control could be responsible for enhancing tuning selectivity. In some instances, when large stimulus patches are used so both the classical receptive field and its surround are stimulated, it is sometimes possible to see a suppressive signal that is tuned in orientation (Fig. 3*C*, right panel), consistent with a role of suppression in enhancing tuning selectivity (Ringach *et al.* 1997, 2002, 2003).

Complex cells and the energy model

Most quantitative models of complex cells derive from the original formulation of Hubel and Wiesel who proposed that they result from the convergence of inputs from a number of simple cells sharing the same preference for orientation (Spitzer & Hochstein, 1985, 1988; see Martinez & Alonso, 2003 for a recent review). One instantiation of this circuit, known as the energy model, considers a pair of linear filters, tuned for orientation and spatial frequency, arranged in quadrature (Adelson & Bergen, 1985). The outputs of the filters are squared and added together to produce a response. The response can be considered a measure of the local signal energy (within a frequency band), therefore the term 'energy model'. Clearly, a LN model applied to a complex cell would not work.

A recent approach used to model complex cells is to learn the input–output function by a two-layer neural network (Lau *et al.* 2002). This method, models the instantaneous rate of firing as

$$\lambda(t) = \sum_{i=1}^{M} \varphi\left(\mathbf{w}_i^T \mathbf{x}\right)$$

Here, $\varphi(\bullet)$ is a sigmoidal non-linearity (a hyperbolic tangent was used in this case). The model is, in effect, an instantiation of the Hubel–Wiesel feed-forward model, as each term can be considered the response of a simple cell. The back-propagation algorithm was used by Lau *et al.* (2002) to minimize the mean-square error on a training dataset and the performance of the model evaluated on a different dataset. The stimulus was a sequence of black/white bars orientated optimally for the cell. Some of the disadvantages of back-propagation are well known: it can settle into local minima and convergence can be rather slow. Nevertheless, the models estimated using this approach, which involves recording the response to flashed bars, were reasonably good at predicting other properties of the neurones, such as the direction selectivity index to a drifting sinusoidal gratings (Lau *et al.* 2002).

A more efficient approach that is yielding interesting results is to study the spike-triggered covariance of the stimulus in response to spatio-temporal Gaussian white-noise. Here, one estimates the covariance matrix $C_{spike} = E\{\mathbf{x}\,\mathbf{x}^T | spike\}$ and compares it to the prior $C_{prior} = E\{\mathbf{x}\,\mathbf{x}^T\}$. Because this matrix is supposed to represent the central second order moment, the result obtained from the spike-triggered average must be subtracted from all the stimuli first (Simoncelli *et al.* 2004). To select directions in stimulus space that appear relevant to establishing the cell's response one computes the eigenvalues of C_{spike} and determines which of these are significantly different from the null distribution of eigenvalues of C_{prior}. Both bootstrap methods (Touryan *et al.* 2002) and analytical results (Everson & Roberts, 2000) can been used to determine the statistical significance of the eigenvalues. Once this is done, the associated eigenvectors provide a subspace of interest that may be further studied by modelling how the neural response depends on the projection of the stimulus

onto the 'relevant subspace'. If an eigenvalue of C_{spike} is significantly larger than expected by chance, the associated eigenvalue is said to lie within the 'excitatory' subspace. Similarly, if an eigenvalue of C_{spike} is significantly lower than expected its eigenvector denotes a direction in stimulus space that suppresses the cell's response. Therefore, the eigenvector is said to lie within the 'inhibitory' subspace.

An example of this method applied to complex cells in cat area 17, from the work of Dan and colleagues, is shown in Fig. 4*A*. In this example, the eigenvectors associated with the two significant (excitatory) eigenvalues are tilted in space–time, as expected from one of the opponent pathway in the energy model of a directional complex cell (Adelson & Bergen, 1985). However, the eigenvectors only provide a basis for the 'relevant subspace' and should not be assigned a particular physical significance, such as that they represent simple-cell inputs to the cell. Both excitatory and inhibitory subspaces were observed in a similar study in macaque V1 by Rust *et al.* (2004) (Fig. 4*B*). In this directional cell, the eigenvectors for the excitatory and inhibitory subspaces preferred opposite directions of motion, suggesting a role for active suppression in the generation of direction selectivity. The results of Rust *et al.* (2004) in complex cells suggest that the quadrature-pair model can be refined in two ways. First, more than a pair of filters may be required to characterize the excitatory subspace. Second, a suppressive subspace appears to be required to appropriately model the responses of complex cells.

Estimating the 'relevant-subspace' with non-Gaussian signals

If both the prior and spike-triggered distributions are Gaussian, the spike-triggered covariance has a nice interpretation in terms of the average information provided by the response about the stimulus (de Ruyter van Steveninck & Bialek, 1988; Chechik *et al.* 2004). Sharpee *et al.* (2003) has proposed a method to extend these ideas to non-Gaussian signals, such as naturalistic image sequences. The model put forward is a Markov chain, $\mathbf{x} \rightarrow P_S\mathbf{x} \rightarrow P(spike)$, where the probability of spiking

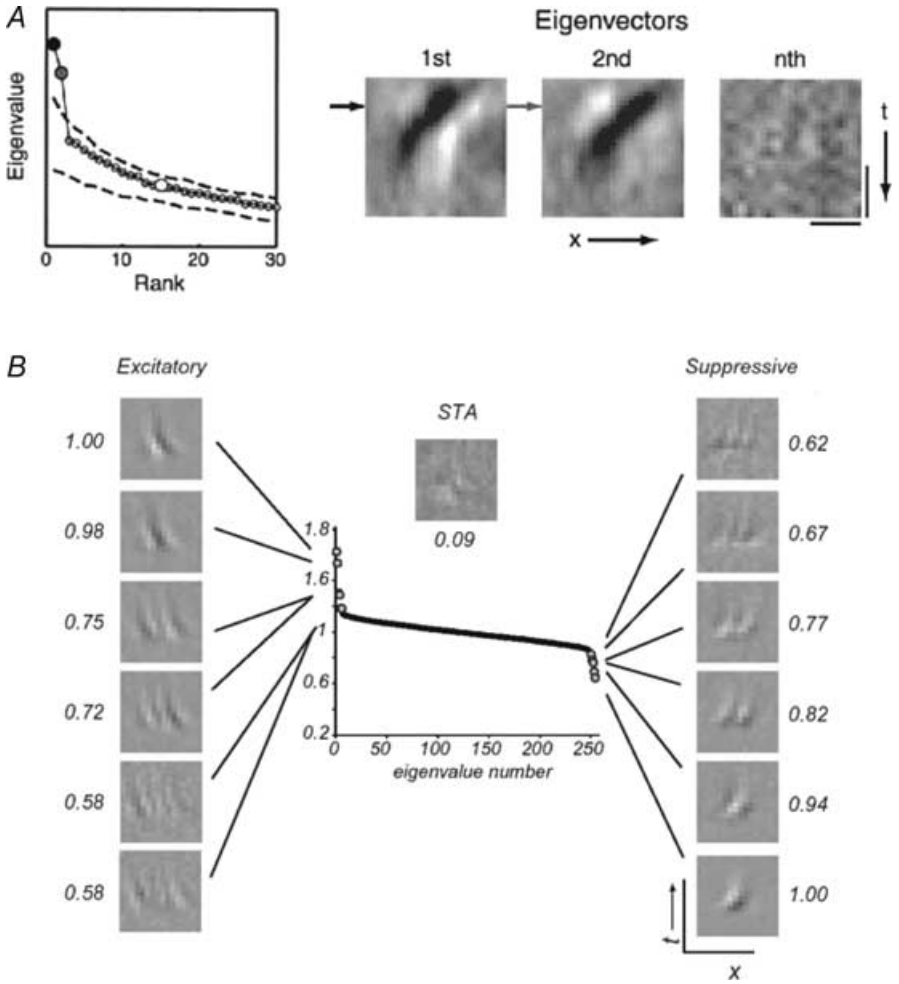

Figure 4. The spike-triggered covariance method
A, results from a complex cell in cat area 17. Only two eigenvalues (excitatory) are significant and the associated eigenvectors (right) show orientated structure in space–time (Touryan *et al.* 2002; copyright 2002 by the Society for Neuroscience). *B*, results from a directional complex cells in monkey V1 (reprinted from Rust *et al.* 2004, copyright 2004, with permission from Elsevier). Both excitatory and inhibitory subspaces can be identified. Eigenvectors within each subspace have a similar orientation in space–time but they have opposite preferences of motion.

depends *solely* on the projection of the input onto a 'relevant subspace', denoted here by $P_S\mathbf{x}$. The subspace is identified by maximizing the mutual information between $P_S\mathbf{x}$ and the neural response. The method may be considered a special case of the information-bottleneck technique (Tishby *et al.* 1999), where the coding of the stimulus is constrained to be linear. The scheme involves the optimization of a function with a large number of parameters, which can be a slow process and it is not guaranteed to converge. The main advantage of the technique is that it can be applied in situations where the signals are non-Gaussian.

Simple/complex cells, the hierarchical model, and theories of cortical function

The original description of simple and complex cells and the associated hierarchical model proposed by Hubel and Wiesel have had a strong impact in shaping theories of V1 function. This framework led many investigators to first develop theories of how simple cells represent the image, deferring the question about the function of complex cells (Maffei & Fiorentini, 1973; De Valois *et al.* 1979; Kulikowski & Bishop, 1981; Olshausen & Field, 1996; Bell & Sejnowski, 1997; Olshausen, 2001; Simoncelli & Olshausen, 2001; Hurri & Hyvarinen, 2003). The hierarchical model has also encouraged the search for coding principles that, when applied layer after layer in a hierarchy, will develop simple and complex-like behaviour (Rao & Ballard, 1997, 1999; Hyvarinen & Hoyer, 2001; Hoyer & Hyvarinen, 2002).

It has been recently suggested, however, that 'simple' and 'complex' cells may represent the ends of a continuum instead of two-discrete classes of neurones (Chance *et al.* 1999; Abbott & Chance, 2002; Mechler & Ringach, 2002). First, it has been demonstrated that the bimodality of the spike modulation ratio (or the F_1/F_0 ratio), taken to validate the existence of discrete classes of neurones (Skottun *et al.* 1991), is likely to be a consequence of the output rectification in what appears to be an otherwise unimodal populations of cells (Mechler & Ringach, 2002; Priebe N, Ferster D, Carandini M & Mechler F, unpublished observations). Second, the distribution of the F_1/F_0 ratio does not show a strong dependence with laminar location as predicted by the hierarchical model. Simple and complex cells (defined according to the F_1/F_0 ratio) are found in all cortical layers, both in monkey (Ringach *et al.* 2002*b*) and cat (Jacob *et al.* 2003). Because the F_1/F_0 ratio is not a direct measure of subfield overlap, it remains possible that a laminar segregation could be determined based on the relationship between 'on/off' subregions (Conway & Livingstone, 2003; Martinez & Alonso, 2003; Hirsch, 2003; Kagan *et al.* 2003), but this remains to be determined.

The discreteness of simple/complex cells is more than a mere technical discussion about how to define these classes of neurones. The question is if the cortex can be considered to be composed of a hierarchy of distinct classes of receptive fields or not. The alternative is that receptive fields lie along a continuum, with simple and complex cells at the ends. A continuum of characteristics appears to hold for other receptive field attributes, such as colour tuning, length summation, spontaneous firing rate, etc. This is not to say that receptive field properties do not *correlate* across the population – they clearly do. As an example, 'simple' cells tend to have lower spontaneous rates than 'complex' cells (see discussion in Mechler & Ringach, 2002). However, the fact that several receptive field properties correlate does not constitute evidence of discreteness.

If we accept the view that receptive field properties appear to lie on a continuum, it would make sense to seek theoretical models that explain the *distribution* of receptive field properties *and their correlations* across the entire population, as well as trends in receptive field properties with laminar location. Such theories would have a quite different flavor from the ones that assume a 'building-block' cortex with simple and complex cells organized in strict hierarchy. Thus, the discreteness of neural populations in the cortex is something we must consider seriously, as the outcome may have a strong impact on how one views cortical organization and function. These are questions that, thanks to advances in receptive field mapping, can now be addressed in a rigorous manner.

Abbott LF & Chance FS (2002). Rethinking the taxonomy of visual neurons. *Nat Neurosci* **5**, 391–392.

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**, 825–837.

Albrecht DG, Geisler WS & Crane AM (2003). Nonlinear properties of visual cortex neurons: Temporal dynamics, stimulus selectivity, neural performance. In *The Visual Neurosciences*, ed. Chalupa L & Werner J, pp. 825–837. MIT Press, Boston.

Anzai A, Ohzawa I & Freeman RD (1999). Neural mechanisms for processing binocular information I. Simple cells. *J Neurophysiol* **82**, 891–908.

Baker CL Jr (2001). Linear filtering and nonlinear interactions in direction-selective visual cortex neurons: a noise correlation analysis. *Vis Neurosci* **18**, 465–485.

Barlow HB (1953). Summation and inhibition in the frog's retina. *J Physiol* **119**, 69–88.

Bell AJ & Sejnowski TJ (1997). The 'independent components' of natural scenes are edge filters. *Vision Res* **37**, 3327–3338.

Benardete EA & Kaplan E (1999). The dynamics of primate M retinal ganglion cells. *Vis Neurosci* **16**, 355–368.

Benardete EA, Kaplan E & Knight BW (1992). Contrast gain control in the primate retina: P cells are not X-like, some M cells are. *Vis Neurosci* **8**, 483–486.

Bonds AB (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Vis Neurosci* **2**, 41–55.

Bonds AB (1991). Temporal dynamics of contrast gain in single cells of cat striate cortex. *Vis Neurosci* **6**, 239–255.

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 & Senn W (2002). A synaptic explanation of suppression in visual cortex. *J Neurosci* **22**, 10053–10065.

Cavanaugh JR, Bair W & Movshon JA (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol* **88**, 2530–2546.

Chance FS, Nelson SB & Abbott LF (1999). Complex cells as cortically amplified simple cells. *Nat Neurosci* **2**, 277–282.

Chechik G, Globerson A, Tishby N & Weiss Y (2004). Information bottleneck for Gaussian variables. In *Advances in Neural Information Processing Systems 16,* ed. Thrun S, Lawrence S, Schölhopf B, MIT Press, Cambridge, MA.

Chichilnisky EJ (2001). A simple white noise analysis of neuronal light responses. *Network* **12**, 199–213.

Conway BR & Livingstone MS (2003). Space-time maps and two–bar interactions of different classes of direction-selective cells in macaque V-1. *J Neurophysiol* **89**, 2726–2742.

DeAngelis GC, Ohzawa I & Freeman RD (1993*a*). 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.

DeAngelis GC, Ohzawa I & Freeman RD (1993*b*). 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 (1995). Receptive-field dynamics in the central visual pathways. *Trends Neurosci* **18**, 451–458.

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.

de Ruyter van Steveninck R & Bialek W (1988). Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. *Proc R Soc Lond B Biol Sci* **234**, 379–414.

De Valois KK, De Valois RL & Yund EW (1979). Responses of striate cortex cells to grating and checkerboard patterns. *J Physiol* **291**, 483–505.

Emerson RC, Bergen JR & Adelson EH (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res* **32**, 203–218.

Enroth-Cugell C & Robson JG (1966). The contrast sensitivity of retinal ganglion cells of the cat. *J Physiol* **187**, 517–552.

Everson R & Roberts S (2000). Inferring the eigenvalues of covariance matrices from limited, noisy data. *IEEE Trans Signal Proc* **48**, 2083.

Geisler WS & Albrecht DG (1991). Cortical neurons: isolation of contrast gain control. *Vision Res* **32**, 1409–1410.

Hartline HK & Ratliff F (1957). Inhibitory interaction of receptor units in the eye of Limulus. *J General Physiol* **40**, 357–376.

Hartline HK & Ratliff F (1958). Spatial summation of inhibitory influences in the eye of Limulus, and the mutual interaction of receptor units. *J General Physiol* **41**, 1049–1066.

Hartline HK, Wagner HG & Ratliff F (1956). Inhibition in the eye of Limulus. *J General Physiol* **39**, 651–673.

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

Heeger DJ (1992). Normalization of cell responses in cat striate cortex. *Vis Neurosci* **9**, 181–197. 92.

Hirsch JA (2003). Synaptic physiology and receptive field structure in the early visual pathway of the cat. *Cereb Cortex* **13**, 1363–1369.

Hochstein S & Shapley RM (1976*a*). Linear and nonlinear spatial subunits in Y cat retinal ganglion cells. *J Physiol* **262**, 265–284.

Hochstein S & Shapley RM (1976*b*). Quantitative analysis of retinal ganglion cell classifications. *J Physiol* **262**, 237–264.

Hoyer PO & Hyvarinen A (2002). A multi-layer sparse coding network learns contour coding from natural images. *Vision Res* **42**, 1593–1605.

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

Hubel DH & Wiesel TN (1968). Receptive fields and functional architecture of monkey striate cortex. *J Physiol* **195**, 215–243.

Hubel DH & Wiesel TN (1977). Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci* **198**, 1–59.

Hubel DH & Wiesel TN (1998). Early exploration of the visual cortex. *Neuron* **20**, 401–412.

Hunter IW & Korenberg MJ (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biol Cybern* **55**, 135–144.

Hurri J & Hyvarinen A (2003). Simple-cell-like receptive fields maximize temporal coherence in natural video. *Neural Comput* **15**, 663–691.

Hyvarinen A & Hoyer PO (2001). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Res* **41**, 2413–2423.

Jacob MS, Peterson MR, Wu A & Freeman RD (2003). Laminar differences in response characteristics of cells in the primary visual cortex. Program No. 910.13. 2003 *Abstract Viewer/Itinerary Planner*. Washington, DC: Society for Neuroscience, 2003. Online.

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 JP & Palmer LA (1987*a*). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol* **58**, 1233–1258.

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

Kagan I, Gur M & Snodderly DM (2003). Spatial organization of receptive fields in V1 neurons of alert monkeys: comparison with responses with gratings. *J Neurophysiol* **88**, 2557–2574.

Kapadia MK, Westheimer G & Gilbert CD (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proc Natl Acad Sci U S A* **96**, 12073–12078.

Kaplan E, Purpura K & Shapley RM (1987). Contrast affects the transmission of visual information through the mammalian lateral geniculate nucleus. *J Physiol* **391**, 267–288.

Kuffler SW (1953). Discharge patterns and functional organization of mammalian retina. *J Neurophysiol* **16**, 37–68.

Kulikowski JJ & Bishop PO (1981). Fourier analysis and spatial representation in the visual cortex. *Experientia* **37**, 160–163.

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

Maffei L & Fiorentini A (1973). The visual cortex as a spatial frequency analyser. *Vision Res* **13**, 1255–1267.

Marmarelis PZ & Marmalelis VZ (1978). *Analysis of Physiological systems: The White-Noise Approach*. Plenum Press, New York.

Martinez LM & Alonso JM (2003). Complex receptive fields in primary visual cortex. *Neuroscientist* **9**, 317–331.

Mechler F & Ringach DL (2002). On the classification of simple and complex cells. *Vision Res* **42**, 1017–1033.

Morrone MC, Burr DC & Maffei L (1982). Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proc R Soc Lond B Biol Sci* **216**, 335–354.

Morrone MC, Burr DC & Speed HD (1987). Cross-orientation inhibition in cat is GABA mediated. *Exp Brain Res* **67**, 635–644.

Movshon JA, Thompson ID & Tolhurst DJ (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J Physiol* **283**, 53–77.

Nykamp DQ & Ringach DL (2002). Full identification of a linear-nonlinear system via cross-correlation analysis. *J Vis* **2**, 1–11.

Olshausen BA (2001). Sparse codes and spikes. In *Probabilistic Models of Perception and Brain Function*. MIT Press, Cambridge, MA.

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

Polat U, Mizobe K, Pettet MW, Kasamatsu T & Norcia AM (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature* **391**, 580–584.

Purpura K, Kaplan E & Shapley RM (1988). Background light and the contrast gain of primate P and M retinal ganglion cells. *Proc Natl Acad Sci U S A Jun* **85**, 4534–4537.

Rao RP & Ballard DH (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Comput* **9**, 721–763.

Rao RP & Ballard DH (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* **2**, 79–87.

Reid RC, Soodak RE & Shapley RM (1987). Linear mechanisms of directional selectivity in simple cells of cat striate cortex. *Proc Natl Acad Sci U S A* **84**, 8740–8744.

Reid RC, Soodak RE & Shapley RM (1991). Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *J Neurophysiol* **66**, 505–529.

Ringach DL (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *J Neurophysiol* **88**, 455–463.

Ringach DL, Bredfeldt CE, Shapley RM & Hawken MJ (2002). Suppression of neural responses to nonoptimal stimuli correlates with tuning selectivity in macaque V1. *J Neurophysiol* **87**, 1018–1027.

Ringach DL, Hawken MJ & Shapley R (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature* **387**, 281–284.

Ringach DL, Hawken MJ & Shapley R (2003). Dynamics of orientation tuning in macaque V1: the role of global and tuned suppression. *J Neurophysiol* **90**, 342–352.

Ringach DL, Shapley RM & Hawken MJ (2002*b*). Orientation selectivity in macaque v1: diversity and laminar dependence. *J Neurosci* **22**, 5639–5651.

Robson JG (1991). Neural coding of contrast in the visual system. *Opt Soc Am Techn Digest System* **17**, 152.

Rodieck RW & Stone J (1965*a*). Analysis of receptive fields of cat retinal ganglion cells. *J Neurophysiol* **28**, 832–849.

Rodieck RW & Stone J (1965*b*). Response of cat retinal ganglion cells to moving visual patterns. *J Neurophysiol* **28**, 819–832.

Rust NC, Schwartz O, Movshon JA & Simoncelli EP (2004). Spike-triggered characterization of excitatory and suppressive stimulus dimensions in monkey V1. *Neurocomputing* **58–60**, 793–799.

Sceniak MP, Ringach DL, Hawken MJ & Shapley R (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nat Neurosci* **2**, 733–739.

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

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.

Shapley R, Enroth-Cugell C, Bonds AB & Kirby A (1972). Gain control in the retina and retinal dynamics. *Nature* **236**, 352–353.

Shapley R, Hawken M & Ringach DL (2003). Dynamics of orientation selectivity in the primary visual cortex and the importance of cortical inhibition. *Neuron* **38**, 689–699.

Shapley RM & Victor JD (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *J Physiol Dec* **285**, 275–298.

Shapley R & Victor JD (1979). The contrast gain control of the cat retina. *Vision Res* **19**, 431–434.

Sharpee T, Rust NC & Bialek W (2004). Analyzing neural resposnes to natural signals: Maximally informative dimensions. *Neural Computation* **16**, 223–250.

Simoncelli EP & Olshausen BA (2001). Natural image statistics and neural representation. *Annu Rev Neurosci* **24**, 1193–1216.

Simoncelli EP, Pillow J, Paninski L & Schwartz O (2004). Characterization of neural responses with stochastic stimuli. In *The Cognitive Neurosciences*, ed. Gazzaniga M, (in press).

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

Spitzer H & Hochstein S (1985). A complex-cell receptive-field model. *J Neurophysiol* **53**, 1266–1286.

Spitzer H & Hochstein S (1988). Complex-cell receptive field models. *Prog Neurobiol* **31**, 285–309.

Sutter E (1975). A revised conception of visual receptive fields based on pseudorandom spatio-temporal pattern stimuli. *Proceedings of the Conference on Testing and Identification of Nonlinear Systems*. California Institute of Technology, Pasadena.

Tishby N, Pereira FC & Bialek W (1999). The information bottleneck Method. *Proceedings of the of the 37th Annual Allerton Conference on Communication, Control and Computing*, http://www.csl.uiuc.edu/allerton/index.html.

Tolhurst DJ & Heeger DJ (1997). 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.