LETTER **Communicated by Bartlett Mel**

A Canonical Neural Circuit for Cortical Nonlinear Operations

Minjoon Kouh
kouh@alum.mit.edu
Tomaso Poggio
tp@ai.mit.edu
Center for Biological and Computational Learning, and McGovern Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

A few distinct cortical operations have been postulated over the past few years, suggested by experimental data on nonlinear neural response across different areas in the cortex. Among these, the energy model proposes the summation of quadrature pairs following a squaring nonlinearity in order to explain phase invariance of complex V1 cells. The divisive normalization model assumes a gain-controlling, divisive inhibition to explain sigmoid-like response profiles within a pool of neurons. A gaussian-like operation hypothesizes a bell-shaped response tuned to a specific, optimal pattern of activation of the presynaptic inputs. A max-like operation assumes the selection and transmission of the most active response among a set of neural inputs. We propose that these distinct neural operations can be computed by the same canonical circuitry, involving divisive normalization and polynomial nonlinearities, for different parameter values within the circuit. Hence, this canonical circuit may provide a unifying framework for several circuit models, such as the divisive normalization and the energy models. As a case in point, we consider a feedforward hierarchical model of the ventral pathway of the primate visual cortex, which is built on a combination of the gaussian-like and max-like operations. We show that when the two operations are approximated by the circuit proposed here, the model is capable of generating selective and invariant neural responses and performing object recognition, in good agreement with neurophysiological data.

1 Introduction

Across the cortex, many different types of neural responses have been observed and subsequently described by different models of neural computations. For example, many cortical neurons are found to produce a sigmoid-like response pattern as a function of certain stimulus parameters like the contrast of an image. Such behavior is often modeled by divisive normalization circuits over a pool of inputs (Carandini & Heeger, 1994; Heeger, 1993).

Some cortical neurons are found to produce strong activity for a certain optimal input pattern and weaker activity for other inputs different from the optimal one (Gallant, Connor, Rakshit, Lewis, & Van Essen, 1996; Gross, Rocha-Miranda, & Bender, 1972; Hubel & Wiesel, 1962; Logothetis, Pauls, & Poggio, 1995; Pasupathy & Connor, 2001; Rauschecker, Tian, & Hauser, 1995; Wilson, Turner, & Laurent, 2004). In some simple cases, like orientation tuning in the primary visual cortex, the observed neural tuning may arise from geometrical arrangement of the receptive fields of the afferent neurons (Hubel & Wiesel, 1962). However, tuning behaviors along multiple nonspatial dimensions cannot be explained by the geometry of the inputs alone. With thousands of excitatory and inhibitory synapses, the input dimensionality of a cortical neuron is usually high, and the neural selectivity may be determined by the pattern of its afferent inputs and not just by the sum of their activations. A gaussian-like template-matching function provides a natural, formal description for the multidimensional tuning behaviors of the cortical neurons.

In addition to the sigmoid-like and gaussian-like behaviors, some cortical neurons are found to be tolerant to appearance-altering transformation of a stimulus. For example, in the visual cortex, neural responses can be tolerant to a certain range of position, size, and rotation changes of an object (Hubel & Wiesel, 1962; Ito, Tamura, Fujita, & Tanaka, 1995; Logothetis et al., 1995). Some neurons are also tolerant to multiple cluttering stimuli within the receptive field, and their responses may be described with a maximum operation (Gawne & Martin, 2002; Lampl, Ferster, Poggio, & Riesenhuber, 2004). The phase invariance of some of the complex cells in V1 may be described by the energy model (Adelson & Bergen, 1985), where phase-canceling (quadrature) pairs of neurons are combined with squaring.

These nonlinear behaviors can manifest themselves along various dimensions within the stimulus space, even for a single neuron. For example, a complex cell in V1 (or a motion-selective cell in MT, or a shape-selective cell in IT) may show a tuned response for a particular edge orientation (or velocity, or a face image), a sigmoidal response for the image contrast, and a translation-invariant response within its receptive field (Gross et al., 1972; Hubel & Wiesel, 1962; Simoncelli & Heeger, 1998).

The goal of this letter is to present a model of a canonical neural circuitry that can generate some of the nonlinear neural response properties found in the physiological data, building on other models that have described different types of nonlinearities. It is helpful to make the distinction between the neural-level operations (which act on the responses of the afferent neurons and produce a nonlinear output) and the stimulus-level behaviors (which are measured from a neuron in typical extracellular physiological experiments, as the sensory stimuli—not the direct inputs to the neuron—are manipulated by the experimenters). The observed nonlinear behaviors of a cortical neuron to the external sensory stimuli are the results of several stages of neural computations and provide only indirect evidence for the

neural-level operations. The proposed neural-level operations and the accompanying circuits are, hence, a hypothesis that needs to be tested with detailed biophysical experiments.

In the following, we first review some of the previous experimental and modeling works on the nonlinear response properties of the cortical neurons. We then point out some similarities and common components found in several well-known neural circuit models and show that a wide range of nonlinear operations can be computed by the same biologically plausible, canonical neural circuitry with different parameter values. Finally, we demonstrate that this circuit, repeated within a hierarchical architecture to provide the tuning and the tolerant behaviors, leads to selective and invariant neural representation that can support object recognition in agreement with neurophysiological data.

1.1 Energy and Divisive Normalization Models. The energy or quadrature model like equation 2.2 in Table 1 postulates that the phase invariance observed in spatial or spatiotemporal orientation-selective neurons may be acquired by computing the "energy," or by half-squaring and summing the responses of the quadrature pairs. Such a construct is motivated by the mathematical identity $\cos^2\theta + \sin^2\theta = 1$ (Adelson & Bergen, 1985).

The energy model is related to the classical proposal by Hubel and Wiesel (1962) for the complex cells in V1, as both models require the pooling of the neurons with the same orientation tuning but with different phase tuning. The energy model, however, specifies a particular polynomial nonlinearity, a squaring operation. Although the energy model has been influential, it is rather unclear whether and how this formulation may be extended to describe other types of invariance.

The divisive normalization model like equation 2.3 has been proposed in various contexts. Reichardt, Poggio, and Hausen (1983) considered divisive normalization for gain control and detection of motion discontinuities by the fly visual system, using forward and recurrent shunting inhibition circuits. A similar mechanism was used to explain contrast-dependent, sigmoid-like neural responses (Carandini & Heeger, 1994; Carandini, Heeger, & Movshon, 1997; Heeger, 1993) and center-surround effect (Cavanaugh, Bair, & Movshon, 2002; Séries, Lorenceau, & Frégnac, 2003) in the primary visual cortex. A divisive normalization scheme was also shown to increase the independence of the neural responses, despite the dependencies in the inputs (Schwartz & Simoncelli, 2001) and to stabilize recurrent neural networks (Chance & Abbott, 2000). Divisive normalization is also likely to constitute an important part of the motion selectivity in V1 and MT (Nowlan & Sejnowski, 1995; Rust, Schwartz, Movshon, & Simoncelli, 2005; Rust, Mante, Simoncelli, & Movshon, 2006) and of the attentional mechanisms (Lee, Itti, Koch, & Braun, 1999; Reynolds, Chelazzi, & Desimone, 1999).

Many of the above formulations of divisive normalization assume a squaring nonlinearity between the contrast of the visual stimuli and the prenormalized responses. Furthermore, this fixed nonlinearity is assumed to be equally strong in the numerator and the denominator (i.e., the same exponent of 2 in equation 2.3).

The energy and the divisive normalization models are well known and more often studied (Schwartz, Pillow, Rust, & Simoncelli, 2006). We therefore focus on the following gaussian-like and max-like nonlinearities, which have been less explored.

1.2 Gaussian-Like Tuning. Especially in the sensory areas of the cortex, many neurons respond strongly to some stimuli but weakly to others, as if they are tuned to a certain optimal pattern of activity of their inputs. For example, some of the neurons in the primary visual cortex show gaussian-like tuning in multiple dimensions, such as orientation, spatial frequency, direction, and velocity (Hubel & Wiesel, 1962). Further along the ventral pathway of the primate visual cortex, some neurons in area V4 show tuned responses to different types of gratings or contour features (Gallant et al., 1996; Pasupathy & Connor, 2001), and inferior temporal neurons are tuned in a bell-shaped way to more complex shapes such as the view of an object (Gross et al., 1972; Logothetis et al., 1995). Some neurons in MT are selective for the motion of a global pattern, displaying a gaussian-like tuning behavior to a set of specific motion components of a moving stimulus (Rust et al., 2006). In other sensory modalities, gaussian-like neural selectivities are also reported: for example, the olfactory neurons in flies (Wilson et al., 2004) and auditory neurons in primates (Rauschecker et al., 1995). In the motor system, the activity of a spinal cord neuron can be regarded as being tuned to a particular pattern of force fields or limb movements (Poggio & Bizzi, 2004). Some neurons in hippocampus, known as place cells, are found to be selective for the spatial position of an animal (O'Keefe, 1976). The tuned response of a neuron can be sharp, broad, sparse, or distributed (Kreiman, 2004), but overall, neural selectivity is believed to be one of the major computational strategies for representing and encoding information in the cortex.

From a theoretical side, it has been demonstrated that radial basis function (RBF) networks using a gaussian kernel can learn effectively from "small" training sets and generalize the input-output mapping to a new set of data (Poggio, 1990; Poggio & Bizzi, 2004). Such an architecture and its generalization properties may be the basis of the stimulus-specific tuning behaviors observed in the cortical networks. Notice that an advantage of the networks with the gaussian-like tuning units (as opposed to a perceptron-like network with the sigmoidal neural units only) is the speed and ease of learning the parameters in the network (Moody & Darken, 1989; Poggio & Girosi, 1989). Learning is even easier for a normalized RBF network because

the synaptic weights from the gaussian units to the output are simply the values of the function at the example points (Girosi, Jones, & Poggio, 1995).

These theoretical considerations, along with the experimental data that find the tuned neural responses within the stimulus space, motivate and suggest a neural-level tuning operation. For conceptual convenience (i.e., the specific mathematical form should not be taken too literally), such an operation may be expressed by the gaussian function, $y(\vec{x}) = e^{-|\vec{x}-\vec{w}|^2/2\sigma^2}$, where a scalar value y corresponds to the strength of the neural response or a monotonic function of it (e.g., the number of spikes or firing rate) for a given input $\vec{x}$, which corresponds to the activity of the presynaptic afferent neurons. In the gaussian function, $\vec{w}$ denotes the optimal input pattern that produces the highest output, and σ determines the sharpness of the tuning.

1.3 Max-Like Selection and Tolerance. In the primary visual cortex, many neurons (probably a subset of the "complex" cells) show tuned, selective responses to different orientations of a bar or Cartesian grating stimulus, while being tolerant to small perturbations in the stimulus location (Hubel & Wiesel, 1962). Such translation-invariance properties are found in other parts of the primate visual cortex (Gallant et al., 1996; Ito et al., 1995; Logothetis et al., 1995; Pasupathy & Connor, 2001). As proposed by Fukushima and others (Fukushima, Miyake, & Ito, 1983; Hubel & Wiesel, 1962; Perrett & Oram, 1993; Riesenhuber & Poggio, 1999), a plausible feedforward mechanism for invariance is to pool from the afferent cells tuned to the transformed versions of a stimulus (e.g., translation, scale, and rotation). For example, a complex cell in the primary visual cortex may pool from the simple cells with the same orientation selectivity but with receptive fields at the neighboring positions, so that its output would still be orientation selective but invariant to the exact spatial locations of the stimuli (Hubel & Wiesel, 1962).

The energy model (Adelson & Bergen, 1985) follows a similar pooling construct for building up translation invariance of complex cells in V1, but its formulation has focused on describing the case where the pairs with opposite phase can be identified (such as the Gabor patches with even and odd phases). Alternatively, in Riesenhuber and Poggio (1999), the maximum operation ($y(\vec{x}) = \max_i(x_i)$, or its soft-max approximation), which selects and transmits the strongest response among a pool of several scalar inputs, was suggested for the pooling. It can serve as a more general neural mechanism for generating tolerances not just to translation, but to other stimulus transformations.[1] For example, by pooling together the afferent

[1]For simple image transformations like translation on a uniform background, a summation operation is sufficient to explain the tolerance properties, but for more difficult cases like clutter, the summation alone is not, as each object in the scene would contribute to the sum and provide a highly distorted output. As a result, the summation operation

neurons tuned to the different views of an object, it may be possible to generate a view-invariant output (Perrett & Oram, 1993; Poggio & Edelman, 1990).

A few recent physiological experiments (Gawne & Martin, 2002; Lampl et al., 2004) have reported that when multiple stimuli are simultaneously presented, some neurons produce a max-like response. These results support the hypothesis of a max-like neural operation acting on the afferent neural responses, although the evidence is indirect since the experiments do not control for the inputs to the recorded neuron.

2 Neural Circuits

While there have been a number of experimental and theoretical studies for the energy and the divisive normalization models, it is not obvious how the cortex may implement the gaussian-like and max-like neural operations. In this section, we propose biologically plausible neural circuits that are capable of computing close approximations to the exact gaussian and maximum, under different parameter values. Furthermore, the proposed circuits can also describe the energy model and the divisive normalization model. We consider only a static approximation of the transduction implemented by neurons. A more detailed description would involve the dynamics of realistic synapses and spiking neurons (Knoblich, Bouvrie, & Poggio, 2007).

The key element in the proposed neural circuit is the divisive normalization of the inputs, which may be accomplished by feedforward or lateral shunting inhibition, as exemplified by the two circuits in Figure 1, respectively (see Carandini & Heeger, 1994; Fukushima et al., 1983; Grossberg, 1973; Heeger, 1993; Reichardt et al., 1983; Serre et al., 2005, and appendix A, but also Chance, Abbott, & Reyes, 2002; Holt & Koch, 1997). Evidence for abundant shunting inhibition in cortex is provided by Borg-Graham, Monier, and Frégnac (1998).

Other key elements in the circuits are the monotonic nonlinearities, acting on the positive-valued neural responses. Denoted by p, q, and r in the figure, they represent and approximate different static strengths of nonlinear, monotonic transfer functions. The energy model and many divisive normalization models (Adelson & Bergen, 1985; Carandini & Heeger, 1994; Rust et al., 2005) have used (half) squaring operations only. However, the

can "hallucinate" or confuse the presence of the target object with the presence of several less optimal objects. In contrast, a max-like operation can provide clutter tolerance more effectively, as well as tolerance to other object transformations (Riesenhuber & Poggio, 1999). One of the earlier proposals for the maximum operation was known as the "maximum amplitude filter" (Taylor, 1964). An average operation, which produces an output smaller than just a simple summation (i.e., a sublinear output like the maximum), may work similarly.

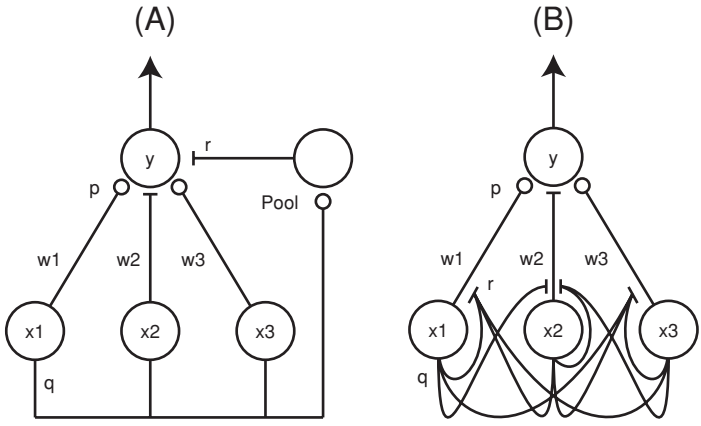


Figure 1: Biologically plausible neural circuits for implementing various neural nonlinear operations in Table 1. The excitatory and inhibitory connections are denoted by an open circle and a bar, respectively. Circuit A performs divisive normalization with feedforward shunting inhibition. The same operation may be computed by lateral inhibition as in circuit B. Such circuits have been proposed and studied earlier in many different contexts (Carandini & Heeger, 1994; Fukushima et al., 1983; Grossberg, 1973; Lee et al., 1999; Nowlan & Sejnowski, 1995; Reichardt et al., 1983; Yu et al., 2002; Yuille & Grzywacz, 1989). Instead of shunting inhibition, the normalization, or gain control, mechanism may also rely on the inherent biophysics of dendritic integration (Borst et al., 1995). Some of the synaptic weights in these circuits may be negative (e.g., w_2), possibly mediated by interneurons (not drawn). Monotonic nonlinearities in the circuits are denoted by p, q, and r. Also see appendix A.

observed nonlinearities are highly variable within the neural population (Britten & Heuer, 1999; Sclar, Maunsell, & Lennie, 1990), and therefore, we assume that the exponents p, q, and r are different and independent in this generalized formulation.[2]

From the biophysical point of view, such different exponents may originate from noise and linear-threshold function between voltage and spike generation (Anderson, Lampl, Gillespie, & Ferster, 2000; Miller & Troyer, 2002), where a wide distribution of polynomial exponents, as large as

[2]Again, care should be taken, as some of referenced studies model the nonlinearity between the stimulus parameter (like the image contrast or the response of a linear filter) and the neural response, whereas the denoted nonlinearities in the proposed circuit are between the pre- and postsynaptic response transduction. Nevertheless, a wide distribution of stimulus-to-response nonlinearity implies that there would be a distribution of nonlinearity in neural response functions as well.

Table 1: Potential Computations That Can Be Performed by the Neural Circuits in Figure 1 at Their Steady States.

Operation	(Steady-State) Output	
Canonical	$y = \dfrac{\sum_{i=1}^{n} w_i \; {x_i}^p}{k + \left(\sum_{i=1}^{n} {x_i}^q\right)^r}$	(2.1)
Energy model	$y = \sum_{i=1}^{2} {x_i}^2$	(2.2)
Sigmoid-like	$y = \dfrac{\sum_{i=1}^{n} {x_i}^2}{k + \sum_{i=1}^{n} {x_i}^2}$	(2.3)
Gaussian-like	$y = \dfrac{\sum_{i=1}^{n} w_i \; x_i}{k + \sum_{i=1}^{n} {x_i}^2}$	(2.4)
Max-like	$y = \dfrac{\sum_{i=1}^{n} {x_i}^3}{k + \sum_{i=1}^{n} {x_i}^2}$	(2.5)

Notes: The canonical form is given in the first row, and its variations are given in the next rows. The responses of the input neurons are denoted by $\vec{x}$ and their synaptic weights by $\vec{w}$. The exponents p, q, and r approximate potential monotonic nonlinear transfer functions in the neural circuit. The summation in the denominator may also include different synaptic weights, as in Fukushima et al. (1983), Rust et al. (2006), and Schwartz and Simoncelli (2001).

four, has been found.[3] Alternatively, different synaptic conductances may generate differential postsynaptic neural activations (Knoblich et al., 2007). For our simulations, the exponents are chosen to be an integer value between 1 and 3, although the general result applies over a wider range.

The combination of two key mechanisms in the circuit (i.e., divisive normalization and polynomial nonlinearity) yields a canonical expression, equation 2.1 (see appendix A for the derivation). Relative strengths of these nonlinear mechanisms can produce various response patterns, as summarized in Table 1. In the case of weak divisive inhibition ($r = 0$)

[3] In Britten and Heuer (1999), using a generalized nonlinear summation model $y = a(\sum_{i=1}^{2} {x_i}^n)^{1/n}) + b$, a similarly wide range of power law exponents, around $n = 2.72$, has been observed. This model can be rewritten as $y = a(\sum_{i=1}^{2} {x_i}^n)/(\sum_{i=1}^{2} {x_i}^n)^{(n-1)/n} + b$, which contains the divisive normalization term more explicitly.

and squaring nonlinearity ($p = 2$), the circuit reduces to the energy model (see equation 2.2), and in the case of squaring nonlinearities ($p = q = 2$ and $r = 1$), the circuit corresponds to the divisive normalization model with matched exponents in the numerator and the denominator (see equation 2.3).

If the strength of nonlinearity in the denominator is stronger than that of the numerator ($p < qr$), y monotonically increases like $|\vec{x}|^p$ for small input, and for large input, y decreases and approaches 0 as the normalizing denominator grows faster than the numerator. In other words, the output y will show a tuning behavior like a gaussian function, peaking at and being selective for a particular pattern of inputs. A simple calculation can show the relationship between the optimal input pattern, synaptic weights $\vec{w}$, and constant k for a given set of parameters p, q, and r (see appendix B). If the strengths of nonlinearity in the numerator and denominator are matched ($p = qr$), then the output y will be a sigmoid-like function (see Figure 2).

For a gaussian function, a Euclidean measure of similarity between two vectors, $|\vec{x} - \vec{w}|^2$, generates a tuned response profile around an optimal input pattern $\vec{w}$. A biologically plausible mechanism for such a computation is suggested by the following mathematical identity that relates the Euclidean distance measure with the normalized dot product:

$$|\vec{x} - \vec{w}|^2 = -2\vec{x} \cdot \vec{w} + 1 + |\vec{w}|^2, \quad \text{if } |\vec{x}| = 1. \tag{2.6}$$

In other words, the similarity between two vectors, $\vec{x}$ and $\vec{w}$, can be measured by either the Euclidean distance or the dot product (the angle between the two vectors in a multidimensional space). Hence, equation 2.6 suggests that a gaussian-like tuning can arise from a normalization and a weighted summation like equation 2.4 (see Maruyama, Girosi, & Poggio, 1992, for more details).

Suppose an additional sigmoid nonlinearity operates on the tuned output y:

$$h(y) = \frac{1}{1 + e^{-\alpha(y-\beta)}}. \tag{2.7}$$

The parameters α and β in equation 2.7 play the similar role as σ of the gaussian function, controlling the sharpness of tuning within the input space for y. These parameters may be learned and calibrated (e.g., large α and β to produce a sharp and sparse tuning), along with the synaptic weights, during the developmental or training period of the neural circuit.

A multiquadric radial basis function, $\sqrt{k^2 + |\vec{x} - \vec{w}|^2}$, has an inverted gaussian-like response profile and may be considered as another possible tuning operation (Girosi et al., 1995). It can be computed with the same

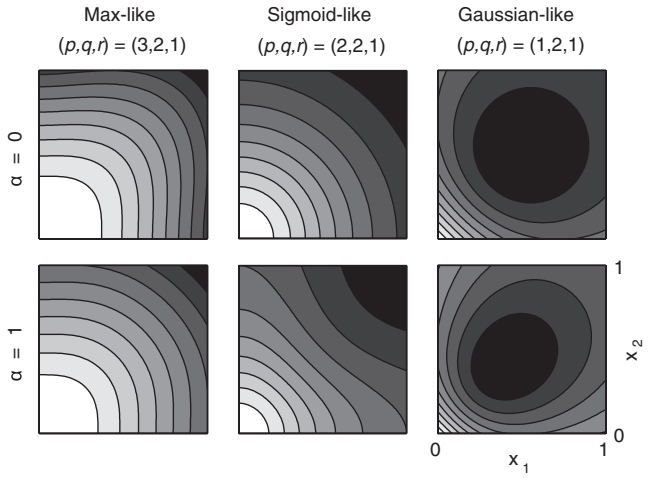


Figure 2: The canonical operation, equation 2.1 in Table 1, possibly computed by the circuits in Figure 1, can generate a wide range of nonlinear neural responses for different sets of parameters p, q, and r, as illustrated in these two-dimensional input spaces. Within each panel, x_1 and x_2 axes represent the responses of two afferent neurons, bounded between 0 and 1, where 1 is the maximum response. The output of the efferent neuron is represented by the gray scale (a darker shade indicates a higher response). The top row shows the default cases. From left to right, the response may be characterized as max-like (equation 2.5), sigmoid-like (equation 2.3), and gaussian-like (equation 2.4), as p goes from 3 to 1 (q and r are fixed at 2 and 1). Similar analyses were done for the responses of the MT neurons in Britten and Heuer (1999). The sigmoid-like neural response in the middle column may also be thought of as a tuned response to the direction of a vector, which is a special case of a gaussian-like response whose tuning peak is located at the far end of the afferent input space. The bottom row illustrates the effect of "tuned" normalization, where each afferent neuron is more strongly normalized by its own response (i.e., $y = \sum_{i=1}^{n} w_i x_i'$, where $x_i' = {x_i}^p/(k + \alpha {x_i}^q + \sum_{j=1}^{n} {x_j}^q)$). As shown here and as considered by Rust et al. (2006), such "tuned" normalization can sharpen the tuning curves within the afferent input space, and it may be implemented by the stronger self-inhibiting synapses in Figure 1B. This mechanism is also related to the proposal of allowing different synaptic weights in the divisive normalization term, as in Schwartz and Simoncelli (2001).

neural circuits in Figure 1 by the gaussian-like tuning followed by an inverted sigmoidal nonlinearity (e.g., $1 - 1/(1 + e^{-\alpha(y-\beta)})$). Although some neurons reportedly show monotonically increasing, multiquadric-like responses along some feature dimensions (Freiwald, Tsao, Tootell, & Livingstone, 2005; Leopold, Bondar, & Giese, 2006), they may still be

explained by the gaussian-like computations operating in the neural input space rather than the stimulus space.

In the limit of weak shunting inhibition, the general computation performed by these canonical circuits can still involve some normalizing behaviors (see Borst, Egelhaaf, & Haag, 1995 and appendix A). Under certain conditions (e.g., negligible normalization due to a large leak current; see appendix A), the operation may be approximated by a weighted summation, still subject to different monotonic nonlinearities (p, q, and r). Excluding some very special cases (e.g., the afferent inputs form opposite-phase quadrature pairs like the energy model, equation 2.2), the overall response of such a circuit increases with the increasing response by any of its afferents (or with the decreasing response by the afferent neurons with the negative synaptic weights). Therefore, the maximum response of the circuit is constrained to be the point with the maximum activations of the positive-weighted afferents and with the minimum activations of the negative-weighted afferents. Tuning to an arbitrary input pattern is not possible, unlike equation 2.4 or a gaussian function.[4]

In contrast to the gaussian-like tuning operation, a max-like output can be achieved if the strength of nonlinearity in the denominator is smaller than the numerator (i.e., $p > qr$), using the same neural circuit and equation 2.5, as proposed by Yu, Giese, and Poggio (2002). In particular, a softmax function ($p = q + 1$, $r = 1$ and small k in equation 2.1) approaches $\max(\vec{x})$ for large q, although our simulation results in the next section indicate that a relatively small value of $q = 2$ may be sufficient to model invariance properties of a neuron.

Therefore, building on the previously proposed neural mechanisms, namely, the distribution of polynomial nonlinearities (Miller & Troyer,

[4]When the number of the afferent neurons is large, however, the weighted summation operation may still create *almost* as complex a tuning behavior as a gaussian function. The afferent response with an inhibitory synaptic weight (e.g., presence of an unpreferred feature in the stimulus) can lower the overall output, acting like a gaussian function tuned to the small input. In other words, the inhibition effectively produces an AND-NOT behavior, and a combination of the AND-NOT operations can create a tuning behavior similar to the gaussian function. The output of the weighted summation operation depends on the weights as well as the norm of the input vector. With a large input dimensionality, the norm of the input vector has less influence over the tuning behavior, as the fluctuations from each input component average out. In other words, the variance of the afferent response decreases with the increasing number of the afferent neurons (cf. $1/n$ dependence of the variance from the n-dimensional statistical samples). As a result, the weighted summation may yield a tuning behavior strongly modulated by the synaptic weights, although it still is a monotonic tuning within the input space of the afferent responses. An approximation to the maximum operation can also be computed by a weighted summation with monotonic nonlinearities: for example, $y = \log(\sum_{i=1}^{n} e^{x_i})$. Similar to the soft-max operation, the summation of strongly scaled input values (e.g., an exponential function) followed by an inverse scaling operation will produce a max-like response. It is an open question whether the biophysics of a real neuron would have enough dynamic range to allow such a steep, exponential nonlinearity.

2002) and divisive normalization (Carandini & Heeger, 1994; Rust et al., 2005; Yu et al., 2002), we have arrived at a generalized nonlinear operation, equation 2.1, which can produce a wide range of nonlinear behaviors as shown in Figure 2. Importantly, all these neural operations may be computed by the same canonical neural circuit, shown in Figure 1, at different parameter values. Furthermore, Figure 3 demonstrates that it is also possible to obtain rather close approximations to the gaussian and maximum operations.

3 Discussion

We have shown that a static approximation of the same neural circuit can generate a wide range of nonlinear neural responses, some of which may be characterized as gaussian-like, max-like, or sigmoid-like. This circuitry can be used to describe other neural models, such as the energy model (Adelson & Bergen, 1985) or the divisive normalization models (Carandini & Heeger, 1994). Therefore, the same biophysical mechanism with different parameters (and potentially even in different operating regimes of the same circuit; Moldakarimov, Rollenhagen, Olson, & Chow, 2005) can compute several types of neural responses.

Such a neural circuit may possibly be operating throughout the cortex. For instance, as shown in Figure 4, a biologically plausible model of the visual cortex, employing the proposed canonical circuitry, can generate shape-selective and transformation-invariant neural response properties that can support object recognition. This model belongs to a class of feedforward, hierarchical models of object recognition (Fukushima et al., 1983; Mel, 1997) and is a much extended version of the original proposal of Riesenhuber and Poggio (1999), based on more detailed anatomy and physiology of the feedforward path of the ventral pathway in the primate visual cortex (see Serre et al., 2005).

Interestingly, the model of motion selectivity in MT by Rust et al. (2006) is equivalent to the proposed neural circuitry in Figure 1, since they both involve weighted summation over the divisively normalized inputs. The model of Rust et al. allows for the negative synaptic weights and for a different set of weights in the denominator (see Figure 2), but these differences are minor (e.g., our current model of object recognition in Figure 4 may be modified to include negative synaptic weights). Possibly, similar neural circuits may also be operating in the dorsal pathway.

The universality of the proposed neural circuit is consistent with two interesting observations: that the basic cortical structure is quite similar within and across different functional areas and that there may exist "canonical microcircuits" in the cortex (Douglas & Martin, 2004; Douglas, Martin, & Whitteridge, 1989; Mountcastle, 2003; Nelson, 2002). The functional distinctions among various neural response properties, hence, may be better considered by a continuous spectrum of nonlinear operations with a canonical

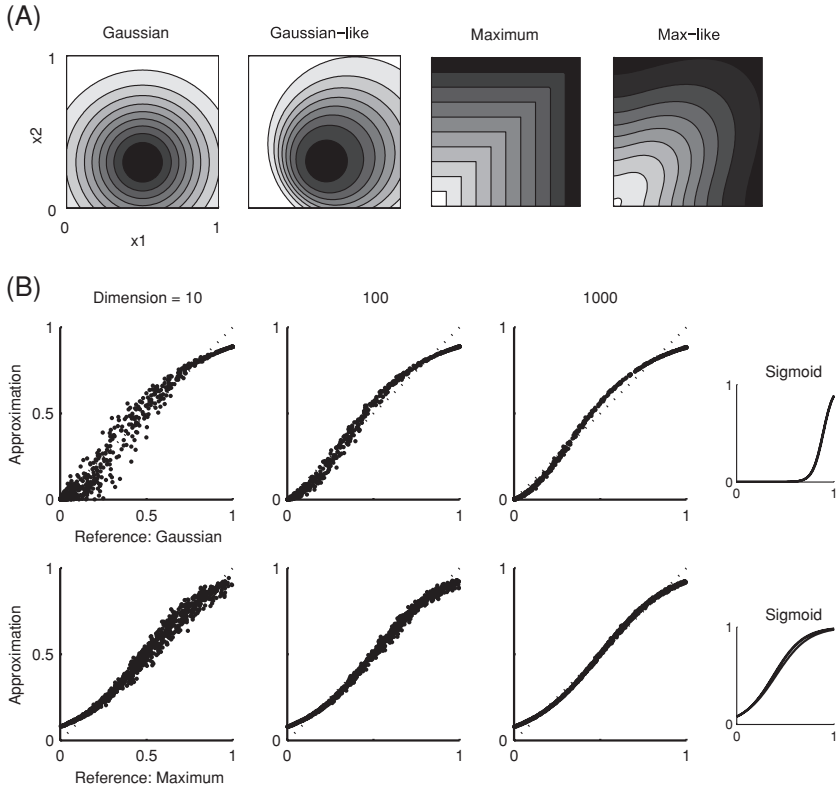
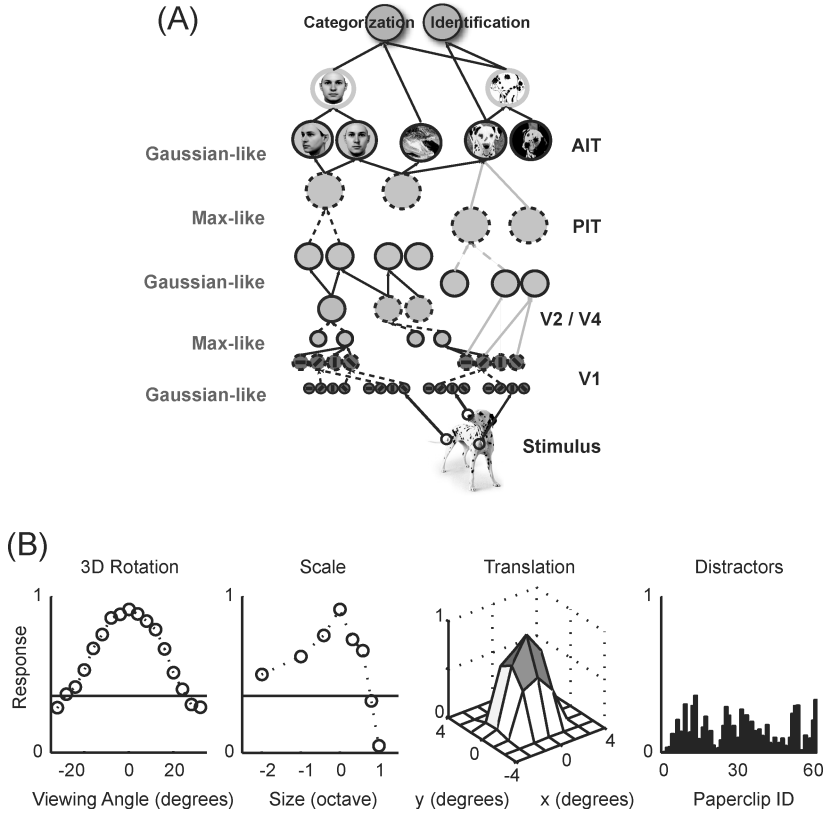


Figure 3: This figure illustrates that the canonical operation, equation 2.1 in Table 1, can approximate the gaussian and maximum operations in two- and higher-dimensional input spaces. The gaussian functions are approximated by equation 2.4 and the maximum operation by equation 2.5, within the afferent input spaces of the indicated dimensionalities ($d = 2, 10, 100$, and 1000). For the gaussian-like operation, the tuning center was found by choosing the synaptic weights $\vec{w}$ and k for fixed $(p, q, r) = (1, 2, 1)$, as shown in appendix B. For the maximum-like operation, we set $(p, q, r) = (3, 2, 1)$ and $\vec{w} = 1$. The sigmoid nonlinearity, equation 2.7, on the output is found by a numerical fitting routine (shown as the insets for the high-dimensional cases). Within each panel of A, the x_1 and x_2 axes represent the responses of two afferent neurons, and the output of the efferent neuron is represented by the gray scale. The reference gaussian is centered at $(0.5, 0.3)$. In each panel of B, the ordinate axis is for equation 2.4 or equation 2.5, and the abscissa is for the reference gaussian or maximum functions; hence, the points along the diagonal line would indicate a perfect fit. These comparisons between the reference and approximating functions are based on 1000 sample points.

circuitry, similar to the proposal that the simple and complex cells in the primary visual cortex may lie on a continuous spectrum of the same neural population (Chance, Nelson, & Abbott, 1999; Mechler & Ringach, 2002; Priebe, Mechler, Carandini, & Ferster, 2004).

Two main mechanisms underlying the operation of equation 2.1 are the divisive normalization and the polynomial nonlinearities. The spontaneous activity or "noise" in the neural circuit plays an important role in generating stable approximations to both mechanisms (Chance & Abbott, 2000; Chance et al., 2002; Miller & Troyer, 2002). The observed variability in the exponents (Britten & Heuer, 1999; Miller & Troyer, 2002) is also important for obtaining a wide range of generalized nonlinear behaviors.

According to the circuit in Figure 1, the neural selectivity ultimately arises from a combination of the feedforward inputs and the intracortical refinements. Such a postulate takes an intermediate position between the two competing theories about the origin of orientation selectivity in the primary visual cortex V1 (Ferster & Miller, 2000). That is, the overall orientation



preference of a V1 neuron is determined by the spatial arrangements of the lateral geniculate nucleus afferent inputs, and then the selectivity may be sharpened by lateral interactions.

As shown by several experimental studies (Douglas et al., 1989; Marino et al., 2005; Wang, Fujita, & Murayama, 2000), inhibitory mechanisms are critically involved in generating neural selectivity and invariance, and in our proposal, the input normalization by a pool cell or lateral inhibition (Figure 1A or 1B) is the major component in the neural circuit. The inhibitory pool cells or interneurons receive input from the afferent neurons with different receptive fields and selectivities. As a result, these inhibitory cells are expected to have a large receptive field (i.e., the union of the receptive fields of the afferent neurons), broad selectivity, and typically high responses.

Figure 4: (A) The architecture of a model of the ventral pathway in the primate visual cortex for object recognition. The model describes the early stages of the visual information processing for "immediate" recognition. Gaussian-like and max-like operations are repeatedly applied in the hierarchy, progressively building up richer and more complex selectivity and invariance properties. See Kouh (2007) and Serre et al. (2005) for a comprehensive description of the current, recently extended version of the model and Riesenhuber and Poggio (1999) for the original proposal. In implementing this model, the same neural operation (equation 2.1 with different parameters; $(p, q, r) = (1, 2, 1)$ or $(3, 2, 1)$ for gaussian-like and max-like operations, respectively) has been used. The software of the model is available at http://cbcl.mit.edu. The layer in the model labeled AIT corresponds to the inferotemporal (IT) cortex, which is the last purely visual area in the hierarchy of the primate visual cortex, and in particular, Logothetis et al. (1995) have found some neurons in anterior IT, tuned to scaled, translated, and rotated (in depth) images of the preferred paperclips that the monkey was trained to recognize. (B) The response of a model IT neuron to the preferred paperclip (under these transformations) and to 60 distractor paperclips is plotted. The maximum distractor response is indicated by the horizontal lines in the first two panels. These tuning curves are obtained by the cascade of the gaussian-, max-, gaussian-, max-, and gaussian-like operations on the stimuli (five layers), and they quantitatively demonstrate that the combination of the canonical operations can achieve view-tuned, selective, and invariant tuning properties. This model unit shows the rotational invariance range of approximately 60 degrees, scale invariance range of over 2 octaves, and translation invariance range of 4 degrees of visual angle, with respect to the maximum distractor responses. The size of the reference object was 2 degrees. Other model units tuned to different paperclips show similar results, in good quantitative agreement with the experimentally observed values. Similar results have been reported earlier with the exact gaussian and maximum operations in a more simplistic model (Riesenhuber & Poggio, 1999).

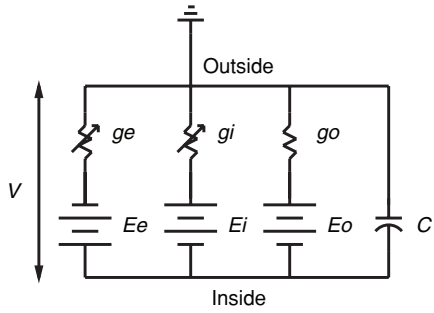


Figure 5: Circuit diagram of the biophysics of a single cell.

Interestingly, putative inhibitory neurons possess similar traits (Bruno & Simons, 2002).

There are other models of cortical nonlinear operations and neural circuits. For example, pattern matching may occur in the dendritic spines (Blackwell, Vogl, & Alkon, 1998). Translation invariance may be computed by the intradendritic mechanisms (Mel, Ruderman, & Archie, 1998) or by the winner-take-all types of neural circuitry (Hahnloser, Sarpeshkar, Mahowald, Douglas, & Seung, 2000; Rousselet, Thorpe, & Fabre-Thorpe, 2003; Yuille & Grzywacz, 1989). The inhibition may operate through presynaptic feedback (Yuille & Grzywacz, 1989) instead of a feedforward pooling mechanism as in Figure 1A. The excitatory and inhibitory circuit elements may be found at specific laminar locations within cortical tissues (Douglas & Martin, 2004; Douglas et al., 1989; Fitzpatrick, 1996). Clearly, the study of more detailed, biophysical implementations of the proposed static neural circuits (e.g., spiking neural network including the kinetics of realistic synapses and of the spike generation process; see section 5 in Serre et al., 2005, and Knoblich et al., 2007) is the next step in the work described here.

Appendix A: Divisive Normalization with Shunting Inhibition

Figure 1A shows a plausible neural circuit that computes a divisive normalization. Figure 5 is an electrical equivalent circuit diagram of the cellular membrane of the output neuron y in Figure 1A, which receives the inputs from the neurons x_1, x_2, and x_3, along with the shunting inhibition from a pool cell. The membrane is assumed to have the capacitance C, the variable conductances g's, and the equilibrium potentials E's. The subscript o (E_o, g_o) is for the leak ionic channels, e (E_e, g_e) is for the excitatory inputs, and i (E_i, g_i) is for the shunting inhibition. The magnitude of the conductances (g_e and g_i) depends on the level of their presynaptic activities.

Suppose that the transformation of a signal from a presynaptic to a postsynaptic neuron is described by the polynomial nonlinearity with the exponent p, q, or r (Miller & Troyer, 2002), as in Figure 1A. Then we may identify

$$g_e \propto \sum_{i=1}^{n} w_i {x_i}^p, \tag{A.1}$$

$$g_i \propto \left(\sum_{i=1}^{n} {x_i}^q \right)^r. \tag{A.2}$$

The membrane potential evolves according to

$$-C\frac{dV}{dt} = g_e(V - E_e) + g_i(V - E_i) + g_o(V - E_o). \tag{A.3}$$

At the steady state,

$$V = \frac{g_e E_e + g_i E_i + g_o E_o}{g_o + g_e + g_i}. \tag{A.4}$$

Now assume that $V_{rest} = E_i = 0$ (i.e., shunting inhibition), $g_e \ll g_i$, based on experimental evidence like Borg-Graham et al. (1998), and assume that E_o is also close to V_{rest}. Then,

$$V \sim \frac{g_e E_e}{g_o + g_i}. \tag{A.5}$$

Next, assume that the membrane potential V and the activity of the output neuron y are approximated by a linear relationship, or a polynomial "nonlinearity" with exponent 1 (Miller & Troyer, 2002) (note that other monotonic relationships can also generate gaussian-like tuning and approximate max-like behaviors). Thus, using equations A.1 and A.2, we arrive at equation 2.1 in Table 1 with positive and negative weights (depolarizing and hyperpolarizing inputs) in general. (For a similar discussion, see Carandini & Heeger, 1994; Reichardt et al., 1983; Chance et al., 2002; Holt & Koch, 1997.)

Note that even without explicitly assuming shunting inhibition (i.e., $g_i = 0$), equation A.4 contains divisive normalization by the input conductance g_e, which can realize similar gain control mechanism (Borst et al., 1995). When g_o is much greater than g_e, the overall operation will be close to the weighted summation without normalization (although g_o is typically found to be on the order of g_e; Borg-Graham et al., 1998).

Appendix B: Optimal Input Pattern for Divisive Normalization

A straightforward calculation on equation 2.1 shows that the peak of the normalized weighted summation operation occurs at

$$x_{oi}{}^{q-p} = \frac{p}{qr} \frac{k + \left(\sum_{j=1}^{n} x_{oj}{}^{q}\right)^{r}}{\left(\sum_{j=1}^{n} w_j \, x_{oj}{}^{p}\right)\left(\sum_{j=1}^{n} x_{oj}{}^{q}\right)^{r-1}} \cdot w_i \tag{B.1}$$

$$\equiv \gamma \cdot w_i. \tag{B.2}$$

The above expression can be interpreted as a tuning around x_{oi} that is proportional to w_i (although due to the exponent $q - p$, the proportionality is not strictly linear).

For consistency, it is required that

$$\gamma = \frac{p}{qr} \frac{k + \gamma^{qr/(q-p)} \left(\sum_{j=1}^{n} w_j{}^{q/(q-p)}\right)^{r}}{\gamma^{p/(q-p)+q(r-1)/(q-p)} \left(\sum_{j=1}^{n} w_j{}^{p/(q-p)+1}\right)\left(\sum_{j=1}^{n} w_j{}^{q/(q-p)}\right)^{r-1}}. \tag{B.3}$$

For a given set of w's, this condition relates the parameters k and γ. If γ is already determined or fixed (e.g., $\gamma = 1$, as in the simulations), equation B.3 will uniquely determine the parameter k. It is unlikely that there is an active neural process that adjusts k for given w's, as the above derivation suggests. Rather, a fixed constant k (possibly originating from leak currents) and a set of synaptic weights $\vec{w}$ determine the center of tuning $\vec{x}_o$.

Note that the relation between k and γ may not always be uniquely defined. Consider a special case when $(p, q, r) = (1, 2, 0.5)$ (divisive normalization by the L2-norm of $\vec{x}$). Then equation B.3 is satisfied for any γ when $k = 0$. This result confirms our intuition that $\vec{w} \cdot \vec{x}/|\vec{x}|$ has the maximum value as long as these two vectors are parallel. However, in general when $p < qr$, equation 2.1 describes tuning around some particular point, $\vec{x}_o$.

Appendix C: Learning the Synaptic Weights

We have shown that both gaussian-like and max-like operations may be implemented by relatively simple, biologically plausible neural circuits. It

is then an interesting question whether these operations can be learned also from a similarly simple and plausible plasticity mechanism, especially for setting the values of the synaptic weights within the neural circuits in Figure 1. According to Table 1, the synaptic weights implementing the max-like operation are all uniform ($w_i = 1$ for all i), indicating that the circuit for the max-like operation probably requires less fine-tuning of the synaptic weights. On the other hand, the gaussian-like tuning operation will likely require more precise tuning of the synaptic weights (Serre et al., 2005).

Hebb's rule ("fire together, wire together") is the best-known synaptic learning rule in neuroscience. Consider the following modified Hebbian rule with a flavor of the stochastic gradient descent algorithm ($\Delta w \propto \Delta y$):

$$w(t + \Delta t) = w(t) + \eta \cdot (y(x, t + \Delta t; w + \eta) - y(x, t; w)), \tag{C.1}$$

where η is a small, random jitter in a synaptic weight. The synaptic weight is strengthened if this random perturbation produces a higher response, and it is weakened if the perturbation produces a lower response. Thus, the "wiring" of the synapse is dependent on not just whether the presynaptic and postsynaptic neurons fire together, but also whether the presynaptic activity has a positive effect on the postsynaptic neuron within a short temporal window (Dan & Poo, 2004) (hence, stochastic gradient descent).[5]

Now suppose that a neural circuit of Figure 1 is repeatedly exposed to a particular pattern of input $\vec{x}_o$, so that a target value for $\vec{w}$ is given. Then the learning rule, equation C.1, brings the synaptic weights to the target after

[5]If the tuning operation does not have a divisive normalization or a gaussian-like response profile, an additional constraint is required for stability. One such example is the following:

$$\sum_{i=1}^{n} w_i^2(t) = w_o. \tag{C.2}$$

This condition, also known as Oja's rule (Oja, 1982), implies that the total synaptic weights are normalized or conserved. When some synaptic weights are increased, others will be decreased. This particular mathematical form (i.e., the squaring of the weights) is chosen for convenience but captures the basic idea. The normalization of the synaptic weights is a separate mechanism, different from the normalization of the inputs. Such a condition is expected since the synapses are likely competing for the limited resources (e.g., the number of the receptors on the postsynaptic membrane). Without any mechanism for competition, a simple Hebb-like rule will be subject to a positive feedback (stronger synapses getting even stronger), leading to instability. There is a growing body of experimental evidence and theoretical arguments in support of the competitive Hebbian mechanism (Miller, 1996). Alternatively, a decay term in the learning rule can also ensure stability, as done in Földiák (1991). In the case of the normalization circuit where the response does not always increase with the larger norm of $\vec{w}$, equation C.1 converges to a stable value without an extra constraint like equation C.2.

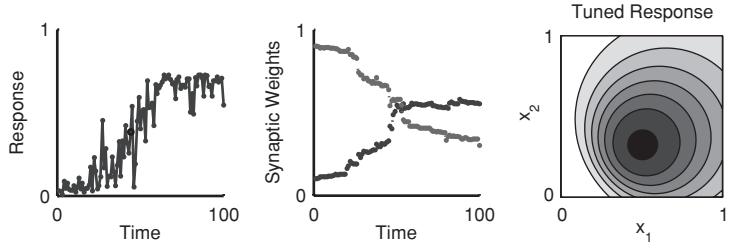


Figure 6: The neural circuit performing the gaussian-like operation can be tuned to a particular input pattern by learning the synaptic weights. The example shown here has two afferent inputs (x_1 and x_2), and the target pattern has been set at $\vec{x}_o = (0.5, 0.3)$. While the presynaptic activations are fixed at $\vec{x}_o$, the synaptic weights evolve according to equation C.1 from randomly selected initial values (middle figure). The left figure shows the evolution of the output of equation 2.4 (with $(p, q, r) = (1, 2, 1)$ and a sigmoid on y), which reaches the maximum as the correct synaptic weights are learned. The right figure shows the learned tuning behavior around $\vec{x}_o$. The random jitters, corresponding to η in equation C.1, were selected from a normal distribution with the maximum value of 0.1.

several iterations (i.e., $\vec{w} \to \vec{w}_o$, so that $y(\vec{x}_o; \vec{w}_o) =$ the maximum output value), as shown by the simulation results in Figure 6.

This scheme involves supervised learning, because the input $\vec{x}$ is required to be fixed at some $\vec{x}_o$ (e.g., the system is exposed to a target stimulus) while the learning takes place. However, choosing the target values themselves may be unsupervised or subject to a hedonistic principle (Seung, 2003). As demonstrated in Serre et al. (2005), the selectivity for the features useful for object recognition (i.e., the centers of the gaussian-like tuning functions) can be learned by storing the snapshots of the naturally occurring sensory stimuli in an unsupervised manner. Such a learning scheme can provide a large, overcomplete set of features or basis functions that reflect the stimulus statistics inherent in the natural environment. We also note that the invariance property can be learned by the similar Hebbian mechanism (Bartlett & Sejnowski, 1998; Földiák, 1991).

Acknowledgments

We thank G. Kreiman, D. Zoccolan, T. Serre, U. Knoblich, H. Jhuang, N. Rust, and M. Riesenhuber, as well as two anonymous reviewers, for helpful discussions.

This letter describes research done at the Center for Biological and Computational Learning, which is in the McGovern Institute for Brain Research at MIT, as well as in the Department of Brain and Cognitive Sciences, which

is affiliated with the Computer Sciences and Artificial Intelligence Laboratory (CSAIL). This research was sponsored by grants from DARPA Contract No. HR0011-04-1-0037, DARPA Contract No. FA8650-06-7632, National Science Foundation-NIH (CRCNS) Contract No. EIA-0218506, and National Institutes of Health (Conte) Contract No. 1 P20 MH66239-01A1. Additional support was provided by Central Research Institute of Electric Power Industry, Daimler-Chrysler AG, Eastman Kodak Company, Honda Research Institute USA, Komatsu, Merrill Lynch, NEC Fund, Oxygen, Siemens Corporate Research, Sony, Sumitomo Metal Industries, Toyota Motor Corporation, and the Eugene McDermott Foundation.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299.

Anderson, J. S., Lampl, I., Gillespie, D. C., & Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science*, *290*(5498), 1968–1972.

Bartlett, M. S., & Sejnowski, T. J. (1998). Learning viewpoint-invariant face representations from visual experience in an attractor network. *Network*, *9*(3), 399–417.

Blackwell, K. T., Vogl, T. P., & Alkon, D. L. (1998). Pattern matching in a model of dendritic spines. *Network*, *9*(1), 107–121.

Borg-Graham, L. J., Monier, C., & Frégnac, Y. (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature*, *393*(6683), 369–373.

Borst, A., Egelhaaf, M., & Haag, J. (1995). Mechanisms of dendritic integration underlying gain control in fly motion-sensitive interneurons. *J. Comput. Neurosci.*, *2*(1), 5–18.

Britten, K. H., & Heuer, H. W. (1999). Spatial summation in the receptive fields of MT neurons. *J. Neurosci.*, *19*(12), 5074–5084.

Bruno, R. M., & Simons, D. J. (2002). Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. *J. Neurosci.*, *22*(24), 10966–10975.

Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, *264*(5163), 1333–1336.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.*, *17*(21), 8621–8644.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J. Neurophysiol.*, *88*(5), 2530–2546.

Chance, F. S., & Abbott, L. F. (2000). Divisive inhibition in recurrent networks. *Network*, *11*(2), 119–129.

Chance, F. S., Abbott, L. F., & Reyes, A. D. (2002). Gain modulation from background synaptic input. *Neuron*, *35*(4), 773–782.

Chance, F. S., Nelson, S. B., & Abbott, L. F. (1999). Complex cells as cortically amplified simple cells. *Nat. Neurosci.*, *2*(3), 277–282.

Dan, Y., & Poo, M. M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, *44*(1), 23–30.

Douglas, R. J., & Martin, K. A. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.*, *27*, 419–451.

Douglas, R. J., Martin, K. A., & Whitteridge, D. (1989). A canonical microcircuit for neocortex. *Neural Comput.*, *1*, 480–488.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.*, *23*, 441–471.

Fitzpatrick, D. (1996). The functional organization of local circuits in visual cortex: Insights from the study of tree shrew striate cortex. *Cereb. Cortex*, *6*(3), 329–341.

Földiák, P. (1991). Learning invariance from transformation sequences. *Neural Comput.*, *3*, 194–200.

Freiwald, W. A., Tsao, D. Y., Tootell, R. B., & Livingstone, M. S. (2005). Single-unit recording in an fMRI-identified macaque face patch. II. Coding along multiple feature axes. In *Society for Neuroscience*, Program No. 362.6. Washington, DC.

Fukushima, K., Miyake, S., & Ito, T. (1983). Neocognitron: A neural network model for a mechanism of visual pattern recognition. *IEEE Transactions on Systems, Man and Cybernetics*, *13*, 826–834.

Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W., & Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J. Neurophysiol.*, *76*(4), 2718–2739.

Gawne, T. J., & Martin, J. M. (2002). Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *J. Neurophysiol.*, *88*(3), 1128–1135.

Girosi, F., Jones, M., & Poggio, T. (1995). Regularization theory and neural networks architectures. *Neural Comput.*, *7*(2), 219–269.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol.*, *35*(1), 96–111.

Grossberg, S. (1973). Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, *52*(3), 213–257.

Hahnloser, R. H., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., & Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*, *405*(6789), 947–951.

Heeger, D. J. (1993). Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J. Neurophysiol.*, *70*(5), 1885–1898.

Holt, G. R., & Koch, C. (1997). Shunting inhibition does not have a divisive effect on firing rates. *Neural Comput.*, *9*(5), 1001–1013.

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

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J. Neurophysiol.*, *73*(1), 218–226.

Knoblich, U., Bouvrie, J., & Poggio, T. (2007). Biophysical models of neural computation: Max and tuning circuits (Tech. Rep. CBCL Paper). Cambridge, MA: MIT.

Kouh, M. (2007). *Toward a more biologically plausible model of object recognition*. Ph.D. Thesis, MIT.

Kreiman, G. (2004). Neural coding: Computational and biophysical perspectives. *Physics of Life Reviews*, *1*(2), 71–102.

Lampl, I., Ferster, D., Poggio, T., & Riesenhuber, M. (2004). Intracellular measurements of spatial integration and the max operation in complex cells of the cat primary visual cortex. *J. Neurophysiol.*, *92*(5), 2704–2713.

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nat. Neurosci.*, *2*(4), 375–381.

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, *442*(7102), 572–575.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.*, *5*(5), 552–563.

Marino, J., Schummers, J., Lyon, D. C., Schwabe, L., Beck, O., Wiesing, P., Obermayer, K., & Sur, M. (2005). Invariant computations in local cortical networks with balanced excitation and inhibition. *Nat. Neurosci.*, *8*(2), 194–201.

Maruyama, M., Girosi, F., & Poggio, T. (1992). A connection between GRBF and MLP (Tech. Rep. AI Memo 1291). Cambridge, MA: MIT.

Mechler, F., & Ringach, D. L. (2002). On the classification of simple and complex cells. *Vision Research*, *42*(8), 1017–1033.

Mel, B. W. (1997). SEEMORE: Combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition. *Neural Comput.*, *9*(4), 777–804.

Mel, B. W., Ruderman, D. L., & Archie, K. A. (1998). Translation-invariant orientation tuning in visual "complex" cells could derive from intradendritic computations. *J. Neurosci.*, *18*(11), 4325–4334.

Miller, K. D. (1996). Synaptic economics: Competition and cooperation in synaptic plasticity. *Neuron*, *17*(3), 371–374.

Miller, K. D., & Troyer, T. W. (2002). Neural noise can explain expansive, power-law nonlinearities in neural response functions. *J. Neurophysiol.*, *87*(2), 653–659.

Moldakarimov, S., Rollenhagen, J. E., Olson, C. R., & Chow, C. C. (2005). Competitive dynamics in cortical responses to visual stimuli. *J. Neurophysiol.*, *94*(5), 3388–3396.

Moody, J., & Darken, C. (1989). Fast learning in networks of locally-tuned processing units. *Neural Comput.*, *1*(2), 289–303.

Mountcastle, V. B. (2003). Introduction: Computation in cortical columns. *Cereb. Cortex*, *13*(1), 2–4.

Nelson, S. (2002). Cortical microcircuits: Diverse or canonical? *Neuron*, *36*(1), 19–27.

Nowlan, S. J., & Sejnowski, T. J. (1995). A selection model for motion processing in area MT of primates. *J. Neurosci.*, *15*(2), 1195–1214.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental Neurology*, *51*(1), 78–109.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J. Math. Biol.*, *15*(3), 267–273.

Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *J. Neurophysiol.*, *86*(5), 2505–2519.

Perrett, D. I., & Oram, M. W. (1993). Neurophysiology of shape processing. *Image and Vision Computing*, *11*(6), 317–333.

Poggio, T. (1990). A theory of how the brain might work. *Cold Spring Harb. Symp. Quant. Biol.*, *55*, 899–910.

Poggio, T., & Bizzi, E. (2004). Generalization in vision and motor control. *Nature*, *431*(7010), 768–774.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, *343*(6255), 263–266.

Poggio, T., & Girosi, F. (1989). A theory of networks for approximation and learning (Tech. Rep. AI Memo 1140). Cambridge, MA: MIT.

Priebe, N. J., Mechler, F., Carandini, M., & Ferster, D. (2004). The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nat. Neurosci.*, *7*(10), 1113–1122.

Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, *268*(5207), 111–114.

Reichardt, W., Poggio, T., & Hausen, K. (1983). Figure-ground discrimination by relative movement in the visual system of the fly. *Biological Cybernetics*, *46*, 1–30.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.*, *19*(5), 1736–1753.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.*, *2*(11), 1019–1025.

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2003). Taking the max from neuronal responses. *Trends. Cogn. Sci.*, *7*(3), 99–102.

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nat. Neurosci.*, *9*(11), 1421–1431.

Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal elements of macaque V1 receptive fields. *Neuron*, *46*(6), 945–956.

Schwartz, O., Pillow, J. W., Rust, N. C., & Simoncelli, E. P. (2006). Spike-triggered neural characterization. *J. Vision*, *6*(4), 484–507.

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nat. Neurosci.*, *4*(8), 819–825.

Sclar, G., Maunsell, J. H., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, *30*(1), 1–10.

Séries, P., Lorenceau, J., & Frégnac, Y. (2003). The "silent" surround of V1 receptive fields: Theory and experiments. *J. Physiol. Paris*, *97*(4–6), 453–474.

Serre, T., Kouh, M., Cadieu, C., Knoblich, U., Kreiman, G., & Poggio, T. (2005). A theory of object recognition: Computations and circuits in the feedforward path of the ventral stream in primate visual cortex (Tech. Rep. AI Memo 036). Cambridge, MA: MIT.

Seung, H. S. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron*, *40*(6), 1063–1073.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761.

Taylor, W. K. (1964). Cortico-thalamic organization and memory. *Proc. R. Soc. Lond. B. Biol. Sci.*, *159*, 466–478.

Wang, Y., Fujita, I., & Murayama, Y. (2000). Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nat. Neurosci.*, *3*(8), 807–813.

Wilson, R. I., Turner, G. C., & Laurent, G. (2004). Transformation of olfactory representations in the drosophila antennal lobe. *Science*, *303*(5656), 366–370.

Yu, A. J., Giese, M. A., & Poggio, T. A. (2002). Biophysiologically plausible implementations of the maximum operation. *Neural Comput.*, *14*(12), 2857–2881.

Yuille, A. L., & Grzywacz, N. M. (1989). A winner-take-all mechanism based on presynaptic inhibition feedback. *Neural Comput.*, *1*, 334–347.

Received February 11, 2007; accepted September 5, 2007.