Neuron
Article

Complex Dynamics of V1 Population Responses Explained by a Simple Gain-Control Model

Yiu Fai Sit,[1] Yuzhi Chen,[2] Wilson S. Geisler,[2] Risto Miikkulainen,[1] and Eyal Seidemann[2,*]
[1]Department of Computer Sciences
[2]Department of Psychology and Center for Perceptual Systems
The University of Texas at Austin, 1 University Station, A8000, Austin, TX 78712, USA
*Correspondence: eyal@mail.cps.utexas.edu
DOI 10.1016/j.neuron.2009.08.041

SUMMARY

To understand sensory encoding and decoding, it is essential to characterize the dynamics of population responses in sensory cortical areas. Using voltage-sensitive dye imaging in awake, fixating monkeys, we obtained complete quantitative measurements of the spatiotemporal dynamics of V1 responses over the entire region activated by small, briefly presented stimuli. The responses exhibit several complex properties: they begin to rise approximately simultaneously over the entire active region, but reach their peak more rapidly at the center. However, at stimulus offset the responses fall simultaneously and at the same rate at all locations. Although response onset depends on stimulus contrast, both the peak spatial profile and the offset dynamics are independent of contrast. We show that these results are consistent with a simple population gain-control model that generalizes earlier single-neuron contrast gain-control models. This model provides valuable insight and is likely to be applicable to other brain areas.

INTRODUCTION

Small visual stimuli elicit neural responses that are distributed over a large area in the primate primary visual cortex (V1; e.g., Hubel and Wiesel, 1974; Grinvald et al., 1994), suggesting that even small stimuli are encoded by a large population of neurons in V1. Furthermore, electrophysiological studies in behaving primates suggest that perception is mediated by a population of neurons rather than by single neurons (Parker and Newsome, 1998; Purushothaman and Bradley, 2005). Thus, to understand the encoding and decoding of visual stimuli in the cortex, it is important to characterize the properties of V1 population responses.

One approach is to estimate population responses from single neuron responses. Single unit recordings in V1 have revealed a number of fundamental properties that ought to contribute to the population responses. First, single neurons have receptive fields with a substantial spatial extent that increases rapidly as a function of retinal eccentricity (Hubel and Wiesel, 1974; Van Essen et al., 1984). Second, the response amplitude of single neurons increases nonlinearly with contrast, typically reaching response saturation at low to modest contrasts (Albrecht and Hamilton, 1982). Third, the tuning of single neurons is typically invariant with contrast, even in the saturated response range (Albrecht and Hamilton, 1982; Albrecht and Geisler, 1991; Heeger, 1991, 1992). Fourth, the latency of the response of single neurons decreases as a function of stimulus contrast (Dean and Tolhurst, 1986; Carandini and Heeger, 1994; Albrecht, 1995). Although these properties are common in V1 neurons, there is a vast heterogeneity among the neurons, and thus it is unclear how these properties are combined and manifested at the population level. In addition, single-unit and multiple-unit studies in V1 have focused mainly on responses at or near the center of activity produced by the stimulus. Responses at locations more peripheral to the center of activity are largely unknown.

Here we provide a complete quantitative description of the real-time spatiotemporal dynamics of V1 population responses to a small, briefly presented (200 ms), localized stationary visual stimulus. Most measurements of response properties in V1 have been performed using drifting stimuli with relatively long durations (several seconds) to approximate a steady-state condition. However, natural saccadic inspection of a visual scene typically produces transient stimulation: 200–300 ms fixations separated by rapid eye movements. In addition, although it is common to analyze cortical responses by their peak responses and latencies (phases) for drifting stimuli, the falling edges of the responses can potentially provide useful information for briefly presented stimuli (Bair et al., 2002). Thus, to fully understand the properties of the population responses under natural conditions, it is important to measure the complete time courses of responses to briefly presented stimuli.

We used voltage-sensitive dye imaging (VSDI; Grinvald and Hildesheim, 2004) in alert, fixating monkeys, to measure population responses in the superficial layers of macaque V1 over an area of approximately 1 cm^2. The imaged area covered the entire region activated by the small local stimulus. We found several unexpected properties that are not obvious from single unit responses. First, the spatial profile of the peak response is independent of stimulus contrast. Second, responses start to rise at all locations approximately at the same time, but rise at a faster rate at the center of activity than at peripheral locations. Third, both the latency and steepness of the rising edge of the response depend on stimulus contrast. Finally, after stimulus offset, the

responses at all locations fall simultaneously and at the same rate, regardless of stimulus contrast. These complex properties illustrate the importance of quantitative characterization of population responses in both space and time.

Next, we considered whether there is a general mechanism that can account for these rich dynamics. To do this, we explored several well-known families of computational models. We find that the observed response properties are inconsistent with simple models having a fixed linear operation followed by a nonlinearity that operates within the individual neuron (e.g., a spike threshold or a refractory effect), and with models that use slow lateral connections to explain the difference in the response dynamics at different cortical locations. To account for the observed properties in both time and space, we propose a simple feedforward population gain-control (PGC) model that generalizes earlier normalization models for single V1 neurons (Albrecht and Geisler, 1991; Heeger, 1991, 1992; Carandini and Heeger, 1994; Carandini et al., 1997). In this model, the temporal dynamics and the gain of the local responses are controlled by population activity in the network rather than by the nonlinear properties of individual neurons or synapses. We simulated the early visual pathway from the retina to V1 by a two-stage PGC model. The model's dynamics closely resemble those in the data.

Responses in the retina and LGN also show evidence of nonlinear contrast gain-control (Shapley and Victor, 1978; Sclar et al., 1990; Kaplan and Benardete, 2001), and thus it is an open question whether a significant part of the nonlinearity in V1 responses is inherited from its input. To address this important question, we used the PGC model to predict how the relative contributions of nonlinearities within layers 2-3 in V1 versus its inputs affect the relationship between stimulus size and V1 response amplitude. The results from a VSDI experiment varying stimulus size were consistent with a PGC model in which most of the nonlinear processing occurs in the first stage, suggesting that the nonlinearities observed in the VSDI responses may be mostly implemented prior to the superficial layers in V1.

In summary, our results illustrate the value of quantitative analysis and computational modeling in testing hypotheses regarding the biophysical and anatomical factors underlying neural population activity. We characterized the spatiotemporal properties of the V1 population responses to small, briefly presented, stimuli that are relevant in natural vision, and found that the response dynamics are complex and unexpected from the results in single unit recordings. Importantly, we show that a simple PGC model can qualitatively account for the complex dynamics, suggesting that gain-control is likely to be a general mechanism contributing to neural dynamics in the brain.

RESULTS

Population Responses to a Gabor Stimulus in V1

We used VSDI to measure V1 population responses to a briefly presented stationary Gabor stimulus while a monkey was performing a fixation task. The goal was to characterize the spatiotemporal dynamics of the population responses over the entire area activated by a small Gabor stimulus. The spatiotemporal dynamics of the fine-scale columnar signals, which are modulated by the orientation of the stimulus, were not measured in the current study and will be addressed in future studies.

Peak Responses: Spatial Distribution and Contrast

The peak responses were taken by averaging the responses over a fixed 100 ms window (160 to 260 ms after stimulus onset; shaded region in Figure 2A). The spatial distribution of the peak responses at two contrasts are shown in Figures 1A and 1B. The distributions are well fitted by two-dimensional Gaussians (Figure 1C). The Gaussians are elongated because of the anisotropic mapping of visual space in V1 (Van Essen et al., 1984; Blasdel and Campbell, 2001; Yang et al., 2007). Importantly, as shown in our previous studies (Chen et al., 2006, 2008), the widths of the Gaussian fits are not significantly different across different contrasts for both the major axis ($\sigma = 2.1$ *mm*) and minor axis ($\sigma = 1.8$ *mm*) (one-way analysis of variance [ANOVA], $p > 0.1$; Figures 1D and 1E). The widths of the spatial distributions are hence largely contrast invariant.

We have previously shown that this spread is not significantly affected by the small variability in the monkey's eye position (Chen et al., 2006). The width of the stimulus ($\sigma = 0.167^{\circ}$) maps to only ~0.5 mm on the cortex through the cortical magnification factor alone (3 mm/deg; obtained in the same animal from retinotopy measurements in another experiment [C.R. Palmer, Y. Chen., and E. Seidemann, 2008, Soc. Neurosci., abstract]). The ~2 mm wide spatial profile is mainly due to the size and scatter of V1 receptive fields (Hubel and Wiesel, 1974; Dow et al., 1981; Van Essen et al., 1984), which dictate the cortical point image (McIlwain, 1986). In addition, some of this widening could reflect significant lateral spread of activity through horizontal connections in V1 (Gilbert and Wiesel, 1979; Rockland and Lund, 1983; Martin and Whitteridge, 1984) and significant contribution from feedback connections (Angelucci et al., 2002). One goal of the modeling component of the current study was to examine the possible contribution of these distinct mechanisms to response spread in V1.

Figure 1F shows the average peak response over a circular region of 0.5 mm radius at the center of the response profile (central region outlined in Figure 1C) as a function of stimulus contrast. Similar to single units, the responses follow a sigmoidal function on a log contrast axis; the solid curve is a Naka-Rushton function fitted to the data ($r^2 = 0.98$).

Overview of the Temporal Response Properties at Different Locations

To analyze the response properties at different locations in V1, the imaging pixels were divided into small bins according to their distances from the center of activity. The image was first divided into 0.5 mm wide concentric annular regions, centered at the peak of the spatial response, with the central region being a disc of 0.5 mm radius. The pixels in the central region had an average distance of 0.25 mm from the center, and the average distance increased by 0.5 mm in each annulus. Due to the anisotropic response profile, we considered only the pixels within a 1 mm wide strip along the major axis of the fitted Gaussian profile. Within this strip the relationship between distance and amplitude was nearly constant. The temporal responses of the pixels within each annulus that were also inside of the strip were averaged to

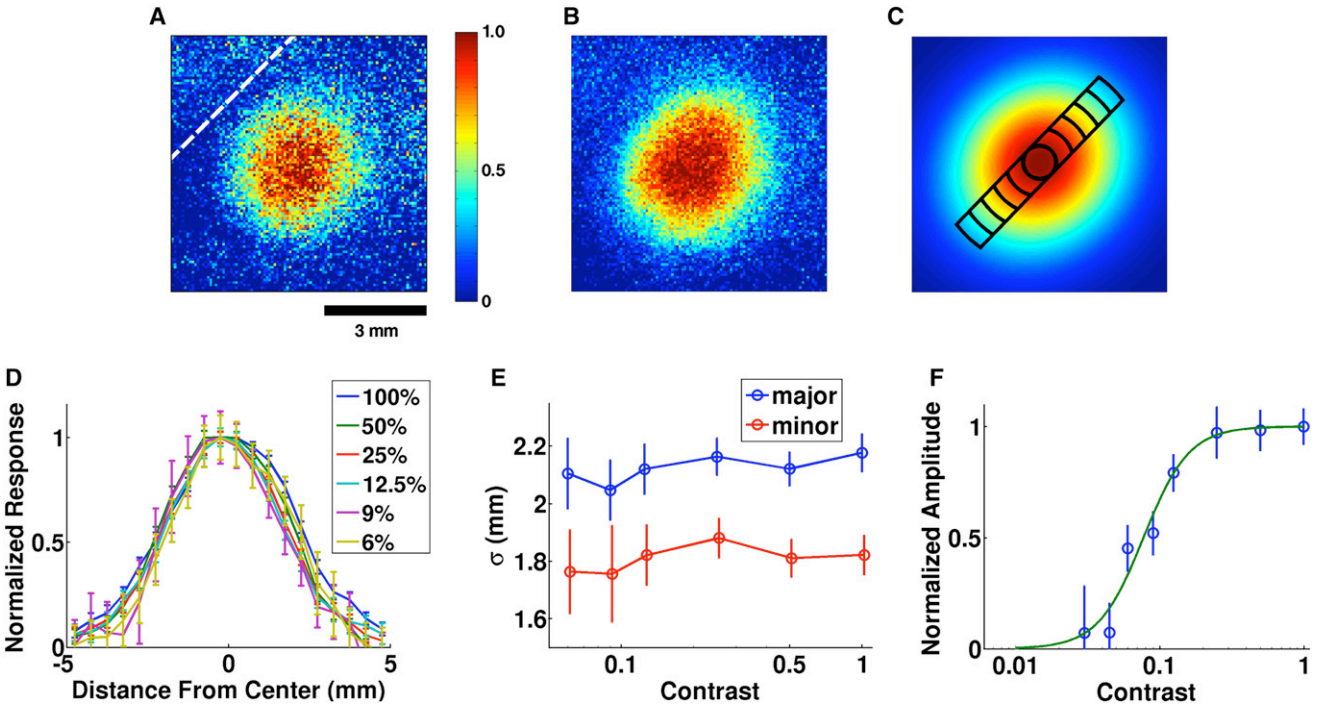


Figure 1. Spatial Distribution and Contrast Response Function of Peak Responses to a Gabor Stimulus

The responses were averaged across all experiments and over a fixed 100 ms time window (shaded region in Figure 2A).

(A and B) Spatial distribution of normalized peak responses to the stimulus at 6% (A) and 100% (B) contrasts. (Dashed white line shows the V1/V2 border.)

(C) 2D Gaussian fit of the response in (B). The outlined regions are the intersection between a 1.0 mm strip along the major axis and six concentric circular annuli of width 0.5 mm. The central annulus is a circle with 0.5 mm radius. The responses of the pixels in the groups that are equidistant from the center are averaged for further analyses in Figures 2 to 4.

(D) Normalized peak responses along the major axis at different contrasts. Here, and in all figures, the error bars are standard errors across individual trials.

(E) The average widths of the Gaussian fits at different contrasts.

(F) Contrast response function at the center. Solid curve is the Naka-Rushton equation ($c^n/(c^n + c_{50}^n)$) fit of the data (open circles, $r^2 = 0.98$). The spatial profiles of the peak responses are therefore largely contrast invariant, even though the responses saturate at high contrast.

produce a single time course for the corresponding distance. Figure 1C shows the bins up to an average distance of 2.75 mm; responses at greater distances were not analyzed because they were weak and noisy, especially at lower contrasts.

Figure 2A shows the average time courses of the responses at the center bin for different stimulus contrasts. To quantitatively characterize these time courses, they were first divided into two parts. The first part, defined as the rising edge, was the response in the first 210 ms after stimulus onset. The rest of the time course was defined as the falling edge. Each individual edge from each trial was smoothed by a five-frame moving average, normalized, and then fitted separately with a logistic function $1/(1+\exp(-\lambda(t-t_{50})))$ (e.g., Figure 2B). The parameter λ describes the slope of the response, and t_{50} is the time that the response reaches half of its peak. For example, a λ of 0.05 means that the response takes about 44 ms after t_{50} to reach 90% of the peak. The same fitting procedure was applied independently at the different locations shown in Figure 1C for each stimulus contrast. We also define the rising edge latency as the time after stimulus onset (t_{10}) for the fitted response to reach 10% of its peak amplitude. Similarly, the latency of a falling edge is the time to decrease by 10% from the peak after stimulus offset. The duration of the rising and falling edges is at least 50 ms (five frames; see Figure 2A), hence there are enough samples for fitting the logistic functions, which provide a good description of the response dynamics.

Figure 2C shows the latencies of the rising and falling edges in the center bin as a function of contrast. As observed in single neuron studies (Dean and Tolhurst, 1986; Carandini and Heeger, 1994; Albrecht, 1995) and Figure 2B, the latency of the rising edge decreases as stimulus contrast increases (one-way ANOVA, $p < 0.01$). On the other hand, there is no significant difference in the falling edge latencies for different contrasts (one-way ANOVA, $p > 0.15$). Asymmetric properties between rising and falling edges are also observed in their slopes (Figure 2D). The slope of the rising edge increases as contrast increases (one-way ANOVA, $p < 0.01$), while the slope of the falling edge remains approximately constant (one-way ANOVA, $p > 0.15$). In other words, the rising edge is accelerated in both latency and slope as contrast increases, while the falling edge is largely independent of contrast.

The space-time color plots in Figure 2E summarize in a compact form all the temporal responses as a function of stimulus contrast and position. The normalized fitted responses for

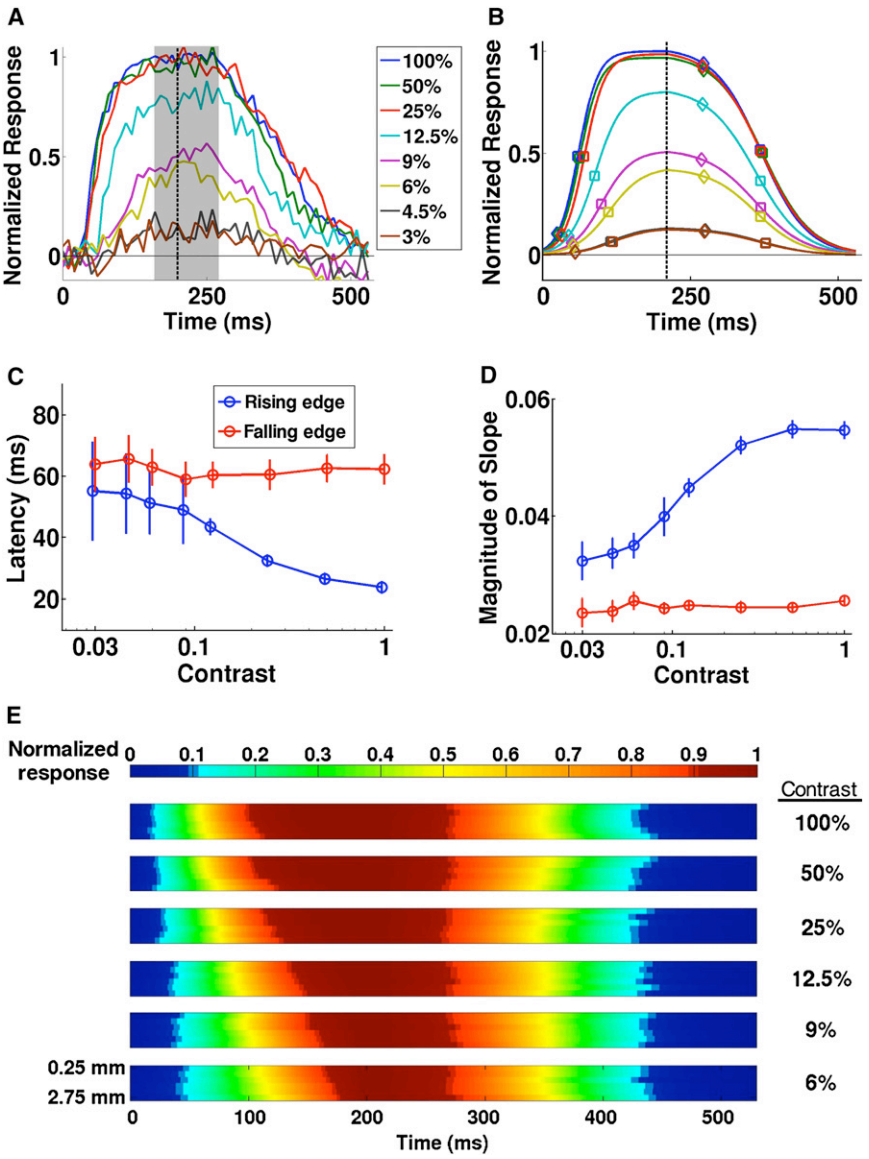


Figure 2. Spatiotemporal Responses to Different Stimulus Contrasts

(A) Time courses of the normalized responses to different contrasts in the center region of Figure 1C. Stimulus was presented at time 0 and disappeared after 200 ms (dotted line). Average responses in the shaded area were used to compute the spatial profiles and contrast response function in Figure 1.

(B) Logistic fits of the time courses in (A). The time courses were divided into two parts by the dashed line. Each part was fitted separately by a logistic function. Diamond and square symbols on each part indicate the latencies (t_{10}) and times to half peak (t_{50}), respectively.

(C and D) Latencies (C) and slopes (D) of the rising and falling edges of the fitted responses as a function of contrast.

(E) Logistic fits of the normalized time courses at different locations for each stimulus contrast. Each horizontal row within the space-time plot for a given contrast shows the fitted time course at one location, from the center (top row) to the outmost region (bottom row). There is a systematic change in slope and latency of the rising edges, whereas the falling edges are similar for different contrasts and locations.

each contrast are shown as separate subplots; within each subplot the time course of the response at each of the six location bins from Figure 1C is indicated by a horizontal row, progressing from the center location at the top to the most peripheral location at the bottom. For example, the upper horizontal row in the plot for 100% contrast corresponds to the dark blue curve in Figure 2B. Several qualitative observations can be made from these maps (they will be quantified later). For each contrast, (1) the response latencies at different locations are approximately equal, as can be seen by the vertically aligned transitions from blue to cyan in each map, and (2) the response rises at a slower rate as distance from the center increases, as can be seen by the increase in the tilt of the transition between the colors as the normalized amplitude increases. In addition, for each location, as contrast increases (3) response latency decreases, and (4) the response rises at a faster rate. Finally, (5) after stimulus offset, the falling edges are similar for all locations and contrasts. All these key properties were also observed in additional experiments in a second animal. We next examine these properties quantitatively.

Properties of the Rising Edges

The above observations can be quantitatively evaluated using the logistic fits obtained from individual trials. Figure 3 plots, for each stimulus contrast, the rising edge latency (t_{10}) as a function of distance from the center bin. For each contrast, there was no significant difference in the latencies at different locations (one-way ANOVA, $p > 0.1$ for all contrasts). However, as observed at the center bin, latency of the rising edge decreased as stimulus contrast increased at all locations (one-way ANOVA, $p < 0.01$ for all locations). These results confirm observations (1) and (3).

Figure 3B plots the same latency data in Figure 3A, but as a function of the peak response. For the same response amplitude, the latency can be different at different stimulus contrasts (e.g., see latencies at a $\Delta F/F = 0.065\%$). This important result demonstrates that the dynamics of the response at a given location in V1 do not depend solely on the local response amplitude, but rather depend on the response amplitudes over a larger region. We will revisit this key property when we discuss possible models of V1 activity.

The rate at which the response rises (λ) depends on both stimulus contrast and on location (Figure 3C). For a particular contrast, the slopes decrease significantly as distance from the

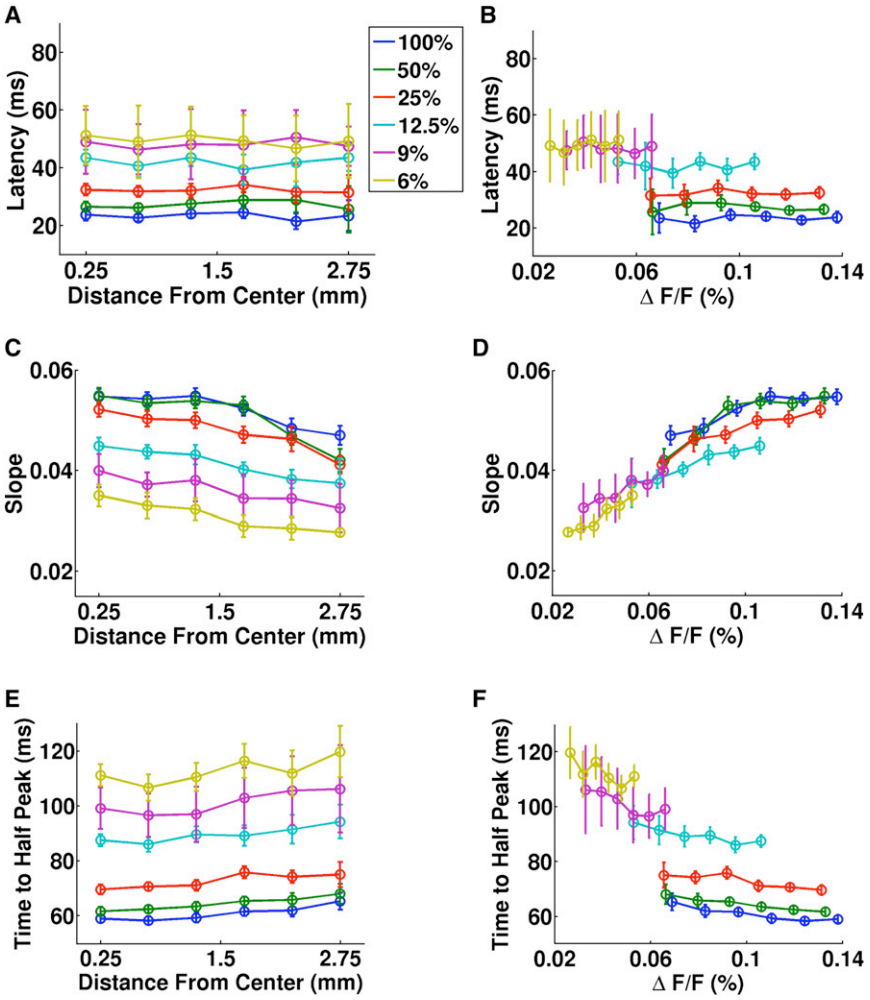


Figure 3. Temporal Properties of the Rising Edges

Time is relative to stimulus onset.

(A) Latencies of the responses at different cortical distances from the center.

(B) Same data as (A), plotted as a function of peak response.

(C and D) Slopes of the responses at different locations (C) and peak responses (D).

(E and F) Time to half of the peak response as a function of location (E) and peak response (F). For a particular contrast, the responses at different locations started to rise at about the same time, but the slopes were shallower at locations that were further away from the center, increasing the times to half peak at these locations.

center increases (one-way ANOVA, $p < 0.01$ for all contrasts), confirming observation (2). In addition, at a fixed location, the slope of the rising edge increases with contrast (one-way ANOVA, $p < 0.01$ for all locations), supporting observation (4). Furthermore, the slope also increases with peak response with a correlation coefficient of 0.94 (Figure 3D).

Due to the decreasing slope as a function of distance from the center, the time to half of the peak response (t_{50}) increased at locations peripheral to the center of activity. If t_{50} was employed as a measure of latency, a traveling wave of activity would appear to be originating from the center (Figures 3E and 3F), as observed previously in anesthetized animals (Grinvald et al., 1994; Jancke et al., 2004; Benucci et al., 2007). The average difference of the time to half peak between the locations 0.25 mm and 2.75 mm away from the center was 6.2 ms. This difference corresponds to a propagation speed of 0.4 mm/ms (for t_{50}), which is at the higher end of the speed of propagation through lateral connections (0.1–0.4 mm/ms; Hirsch and Gilbert, 1991; Murakoshi et al., 1993; Grinvald et al., 1994; Nelson and Katz, 1995; Gonzalez-Burgos et al., 2000). As we shall see later, such differences in time to half of the peak can be explained by a feedforward PGC model.

Properties of the Falling Edges

The dynamics of the falling edges of the responses differed markedly from those of the rising edges. As shown in Figure 4, both the latency and slope were independent of contrast and location (two-way ANOVA, $p > 0.15$ across contrasts and locations for both latency and slope). The responses at all locations therefore fell approximately simultaneously and at the same rate, regardless of the stimulus contrast and response amplitude, supporting observation (5). The latency was slightly longer (t_{10} = 65 ms) and the slope shallower (λ = 0.026) than those of the rising edges. Such asymmetry in the temporal properties of the falling and rising edge can also be explained by the PGC model.

Although the falling edge latency is longer than the rising edge for all contrasts, the reverse relationship has been observed in the firing rates of single units (Bair et al., 2002). The apparent discrepancy is consistent with the fact that the VSDI measures membrane potentials (Grinvald and Hildesheim, 2004). Because of spike threshold, the onset of spiking activity will lag behind the rise in the VSDI response. On the other hand, for the falling edge the drop in spiking activity will coincide with the drop in the VSDI response (as long as membrane potential is above threshold). Thus, for spikes it is quite possible for the onset latency to be greater than the offset latency. Consistent with this possibility, we recently found that the threshold for observing significant spiking activity in V1 to be about 30%–40% of the maximal VSDI response (C.R. Palmer, Y. Chen., and E. Seidemann, 2008, Soc. Neurosci., abstract).

Plausible Families of Models

Is there a simple functional model that can account for these complex spatiotemporal dynamics of V1 population responses? An obvious starting point is to consider previous models that have been proposed to account for the response properties of single neurons. Here we consider three families of such models,

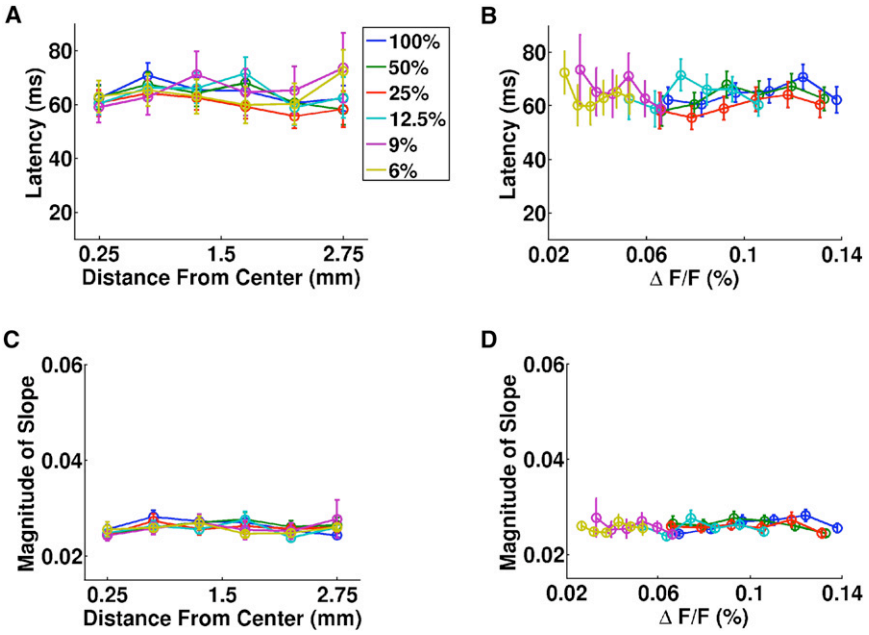


Figure 4. Falling Edge Latencies and Slopes
Time is relative to stimulus offset.
(A and B) Latencies of the responses as a function of location (A) and peak response (B).
(C and D) Absolute value of the slope as a function of location (C) and peak response (D). The latencies and slopes were independent of contrast and location.

scaling them up for population responses by regarding each VSDI pixel as a "single unit" that has the average properties of the neural population falling under that pixel.

Static Nonlinearity Models

A well-known model of single neuron responses is the LN model, which consists of two components: an initial linear weighting function or linear filter (L) followed by a static nonlinearity (N) such as spike threshold or response saturation due to refractory effects. This model is popular because it is relatively simple and easy to analyze. For example, when the model is valid, the receptive field estimated using spike-triggered averaging will equal the initial linear weighting function. Unfortunately, this simple class of models is inconsistent with several key properties of the spatiotemporal dynamics of V1 population responses. First, these models predict that the spatial profile of the responses should widen and change shape as stimulus contrast is increased. This is inconsistent with the observed spatial profiles (Figure 1E). Second, these models predict a longer latency for the falling edge of the response at high contrast, which is not observed in the VSDI responses (see Figure 4). Finally, in an LN model the dynamics of the response are tied to response amplitude, yet V1 responses show clear decoupling of amplitude and latency (Figure 3B). For simulations of the LN model's predictions, see Supplementary Materials 1 (available online).

Lateral Propagation Models

The time to half peak (t_{50}) of the rising edges in the VSDI responses increases with increasing distance from the center of the active region. A common view about such delay is that it results from propagation through slow lateral connections. If lateral spread was indeed the only source of the response beyond a critical distance from the center, then beyond this critical distance the response latency (t_{10}) should increase linearly as a function of distance, which should be evident because of their relatively slow propagation speed (Hirsch and Gilbert, 1991; Murakoshi et al., 1993; Grinvald et al., 1994; Nelson and Katz, 1995; Gonzalez-Burgos et al., 2000). However, there was no significant difference in the rising edge latency of the VSDI responses for a wide range of distances (Figure 3A). In fact, for 50% and 100% contrast stimuli, rising edge latencies remained the same up to a distance of 3.25 mm (within which a reasonably good fit to the data could be obtained; data not shown). These results and additional simulations (see Supplementary Materials 2) suggest that within the examined range the observed dynamics are not consistent with a significant contribution of slow lateral connections to the observed response spread. Below we show that the simple feedforward PGC model can account for the observed dynamics.

Normalization Models

Normalization gain-control models have been used to account for many nonlinear properties of single unit responses in the LGN and V1 (Albrecht and Geisler, 1991; Heeger, 1991, 1992; Carandini and Heeger, 1994; Carandini et al., 1997; Mante et al., 2008). In particular, this family of models can explain response saturation (Albrecht and Hamilton, 1982), contrast-invariant tuning (Sclar and Freeman, 1982; Skottun et al., 1987; Albrecht and Geisler, 1991), and phase advance of response at high stimulus contrasts (Carandini et al., 1997). As we have seen, these properties are also observed in the VSDI responses. Thus, normalization models appear to be more promising than the other two families of models. However, an important question is whether normalization models can account for the other properties observed in the VSDI data, especially the changes in the rising edge at different locations and the invariance of the slope and latency of the falling edge. In the next two sections, we show that a generalization of the normalization model can qualitatively account for all the spatiotemporal properties of the VSDI responses.

Population Gain-Control Model

The PGC model is a generalization of earlier single-neuron normalization models (Albrecht and Geisler, 1991; Heeger, 1991, 1992; Carandini and Heeger, 1994; Carandini et al., 1997; Mante et al., 2008). In contrast to these early models, the PGC model aims at explaining the responses of the *entire* active neural population in V1 in both time and space.

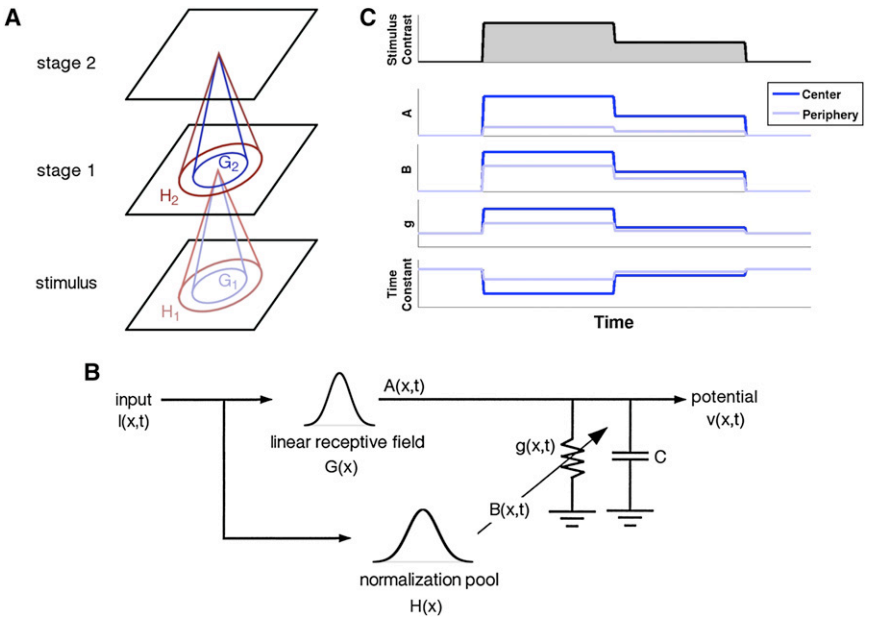


Figure 5. Canonical Feedforward Population Gain-Control Model for the Early Visual Pathway

(A) The model consists of an input layer and two stages. Each layer is modeled by a spatially organized array of units with the same set of parameters. The input layer represents the visual stimulus only and does not employ the model in (B). The processing in the first stage represents the nonlinearity in the retina, LGN, and layer 4 of V1, whose spiking responses are fed into the second stage as input. The second stage represents the superficial layers in V1 where the VSDI signals are measured. The blue and red projections illustrate the receptive fields and normalization pools, respectively, for two example units in the first (light colors) and second (dark colors) stages of the model.

(B) The processing in a model unit. Each unit first computes two weighted sums (with spatial pooling functions $G(x)$ and $H(x)$) which feed into a parallel resistor-capacitor (RC) circuit. The voltage across the capacitor is the unit's response. The conductance $g(x,t)$ of the resistor is inversely related to the gain of the circuit, i.e., it has a divisive effect on the weighted sum $A(x,t)$. It also affects the dynamics of the response. The conductance depends on the weighted sum $B(x,t)$ obtained with normalization spatial pooling function $H(x)$. (Figure modified from Carandini et al., 1997.)

(C) Example time courses of the model's components at the center and periphery in response to a hypothetical stimulus.

In describing the PGC model, we focus on determining the predicted responses to simple Gabor patches like those used in the present experiments. To further simplify the discussion we describe a one-dimensional version of the model, which represents the collapsed data along the major axis (x axis) of the VSDI response profile (see the black rectangular region in Figure 1C). Each layer in the model therefore contains an array of units indexed by x, where each unit represents the average activity of the small neural population within a pixel in the VSDI image. Extension to a full two-dimensional model is straightforward but will not be discussed here.

The visual pathway is represented by a network consisting of an input layer and two stages in Figure 5A. Because VSDI responses in V1 are largely determined by the contrast of the Gabor stimulus, independent of its specific orientation and phase, the stimulus is represented in the input layer by its Gaussian contrast envelope, with magnitude directly proportional to the stimulus contrast. A fixed time delay is also added.

The first stage in the model represents the nonlinear processing that occurs in the retina, LGN, and layer 4 in V1. Although it would be more realistic to model each of these areas individually, there are not enough experimental data at the population level to provide sufficient constraints. The second stage represents layers 2-3 in V1 where the VSDI signals are measured. The units within each stage are identical and implement the filtering and normalization circuit illustrated in Figure 5B.

Processing in the Model Units

At each stage of the model there is an initial step that represents receptive field summation and normalization pooling. The spatial receptive field of each unit x has a Gaussian weight profile $G(x)$ centered at its location, and the spatial normalization pool has a Gaussian weight profile $H(x)$ centered at the same location. The result of the receptive field summation step, $A(x,t)$, then passes through a resistor-capacitor (RC) circuit, whose conductance is controlled by a normalization pool (Figure 5B; Carandini et al., 1997; Mante et al., 2008). As Carandini et al. point out, in such a model the conductance will affect both the gain and the response latency. The voltage across the capacitor, $V(x,t)$, which represents the membrane potential, is the response of the unit.

The key property of the model is that the conductance, $g(x,t)>0$, of the resistor at each unit is not fixed but increases from a baseline value as a function of the weighted average over a local region of the input. For a static current, i.e., $A(x,t)=A(x)$, the steady state of the voltage across the capacitor is $V(x)=A(x)/g(x)$. In other words, the gain of the circuit is the inverse of conductance, and therefore the conductance has a divisive (normalizing) effect on the output of the receptive field summation. The input region that contributes to $g(x,t)$ is called the normalization pool. The overall strength of normalization activity is controlled by a scale factor on the output of the normalization pool.

Response Transformation between Stages

The response at each model pixel represents membrane potential, which dominates the VSDI responses. Because neurons communicate through spikes, the responses in the first stage must be converted into spikes that the second stage receives. In a recent study we found that the VSDI responses are related to spiking activities by a power function (C.R. Palmer, Y. Chen., and E. Seidemann, 2008, Soc. Neurosci., abstract). A similar relationship has been found between average membrane potential and spike rate in single unit recordings (Anderson

et al., 2000; Finn et al., 2007). A fixed power function is thus applied to the responses in the first stage and the results are fed into the second stage as inputs.

Although feeding the stimulus represented in the input layer directly into the first stage provides a reasonably good fit to the data, the predictions shown here are for a model with a fixed power function applied to the activity in the input layer. This initial nonlinearity is plausible given the accelerating point nonlinearities seen in the earliest levels of the visual system; e.g., the nonlinear relationship between the membrane potential of the photoreceptors and their rate of glutamate release (Witkovsky et al., 1997). The model's general behavior is not affected by this nonlinearity.

General Behavior of the Model

One way to understand the dynamics of the model is through the time constant of the RC circuit of each unit (Figure 5C). Note that the time constant and the gain of the circuit are inversely related to conductance. When the normalization activity in a unit is high, the conductance is large and the time constant and gain are small. This property can account for much of the dynamics observed in the VSDI responses. At a particular unit, when the stimulus contrast is high, the receptive field summation and hence the normalization activity is large in all the neighboring units, resulting in faster dynamics (Figure 5C). This property is consistent with the observed dynamics in the rising edges of the VSDI responses. In addition, for a Gabor stimulus, the normalization activity is largest at the center, where the contrast is the highest (Figure 5C). The response at the center therefore rises at a faster rate than those at the periphery, which again is consistent with our observations. This is an interesting property because the spatial difference in gains can account for the traveling wave of the time to half peak observed in the rising edges, which are generally attributed to slow lateral connections.

When there is no input, conductance is at the baseline value in all the units; hence, the temporal dynamics are the same everywhere in the model (Figure 5C). Thus, after input offset, the responses at all the units start to decay at the same time and at the same rate, as observed in the falling edges of the VSDI responses. Finally, the divisive effect of conductance in the model also causes the steady-state response to saturate when the input amplitude is large. The dynamic nonlinearity in the model therefore can account for many of the observed properties of the responses. Interestingly, there is some evidence that the model's prediction of asymmetrical effect of contrast on the rising and falling edges of the response holds approximately for single neurons in the primary visual cortex of cat (see Supplementary Materials 3).

Effects of Normalization Pool Size

How does the width of the normalization pool weighting function affect the rising edge and the spatial profile of the response? Consider a fixed localized input and different Gaussian normalization weighting functions that have the same total weight. If the pool is wide, then the normalization activity will be similar for units near and far from the center of activity. Thus, the difference in the slopes of the rising edges across space will be small. On the other hand, if the normalization pool size is small, there will be a large difference in the time constants of different units. These considerations suggest that the observed difference between the time courses at the different locations could be explained by a feedforward PGC model with an appropriate pool size.

Normalization pool size also influences the spatial profile of the response. Consider a static Gaussian input and its corresponding steady state response, $V(x) = A(x)/g(x)$. If the width of the pool is much wider than the input, then the normalization activity and hence the conductance $g(x)$ will be the same at all units. In this case, the spatial response profile will simply be a scaled version of the receptive field summation, which is a Gaussian. On the other hand, if the normalization pool is much smaller than the input, then response saturation will occur at a different stimulus contrast for each unit, as in the LN model, thus flattening the response profile at high contrasts (Supplementary Materials 1). As a result, to achieve the contrast-invariant spatial profile observed in the VSDI responses, the normalization pool size must be at least comparable to the size of the stimulus.

In sum, the normalization pool size affects both spatial and temporal properties of the responses. Based on these properties, it is possible to estimate the overall pool size from the data analytically (see Supplementary Materials 4).

Simulation of VSDI Responses

VSDI responses were simulated with a network that consisted of an input layer and two subsequent stages (see Figure 5A). Although each stage has its own set of parameters, some of the key parameters were constrained by previous anatomical and physiological measurements reported in the literature (see Experimental Procedures). Figure 6A plots the spatial profiles of the peak responses in the model for different input contrasts. Consistent with the VSDI responses, the widths of the profiles are all the same. Note that profiles will only be contrast invariant in the model for stimuli with sizes that are smaller or comparable to the receptive field of the V1 units; profiles for large stimuli will change shapes and widths as a function of contrast, due to saturation. The contrast response function of the model is plotted in Figure 6B, which provides a good fit to the data ($r^2 = 0.98$).

Figure 6C shows the space-time plot of the predicted VSDI responses. The model captures qualitatively the observed spatiotemporal properties of the responses. For each contrast, (1) the rising edge latencies (t_{10}) at different locations are similar, with a maximum difference of 2 ms, and (2) the slope of the rising edge becomes shallower as distance from the center increases. For each location, as contrast increases, (3) response latency decreases, and (4) the rising edge becomes steeper. Finally, for all contrasts and locations, (5) latencies and slopes of the falling edges are similar (<3 ms difference).

Relative Normalization Strengths in the Different Stages

In the mammalian visual system, contrast gain-control (normalization) has been observed in the retina, LGN, and visual cortex, with a spatial scale that progressively increases (see Introduction). Similarly, in the PGC model, normalization operates at two stages, with the sizes of the receptive fields and normalization pools in the second stage being twice those in the first (Sceniak et al., 2006). An important question is whether our

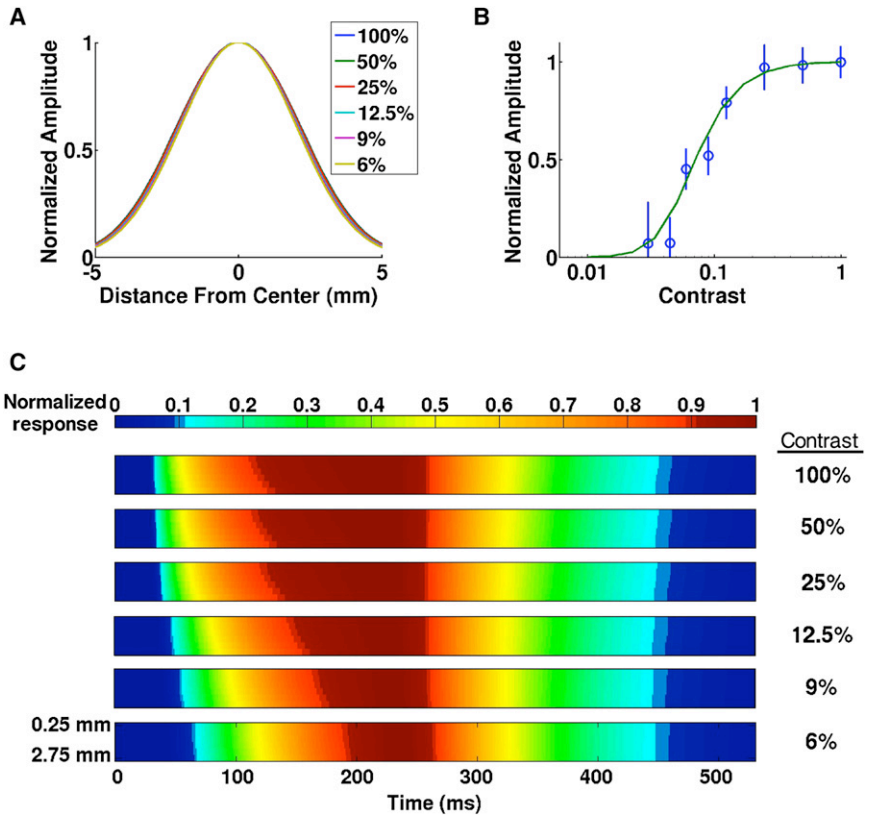


Figure 6. Responses in the Second Stage of the Model

(A) Normalized spatial profile of the peak responses.

(B) Normalized contrast response function at the center. Circles are the responses from the data ($r^2 = 0.98$).

(C) The space-time plot of the normalized responses is qualitatively similar to that of the data (Figure 2E), suggesting that population gain-control may be a general mechanism of visual processing.

results provide evidence concerning the relative strength of the normalization in the different stages of visual processing. As it turns out, the stimuli used in our main experiment are not sufficient to discriminate between hypotheses concerning relative normalization strength.

However, by exploring the PGC model, we found that the relative strength of normalization at the two stages has a large impact on the expected size tuning of V1 responses. Therefore, by varying the size of the stimulus, it is possible to estimate the relative contributions of normalization in the first stage (retina to layer 4 of V1) and the second stage (superficial layers of V1). Figure 7A shows predicted response amplitude at the center of the activated region in the superficial layers of V1 as a function of stimulus size for a 100% contrast Gabor stimulus. Each curve in the figure is for a different strength of normalization in the first stage of the model relative to the total strength in both stages. When normalization only occurs in the first stage of the model (1.00), the response increases with stimulus size because the second stage is linear (i.e., no normalization). As the normalization in the second stage becomes stronger (the other curves), the relative response to the larger stimuli (e.g., $\sigma = 1^\circ$) decreases, because normalization has a divisive effect on the input from the first stage.

In an additional experiment we measured the VSDI responses to 100% contrast Gabor stimuli with $\sigma = 0.167^\circ$ and 1°. The red dots in Figure 7A plot the relative responses to the two stimuli. The peak response to the large stimulus is about 7% less than to the small stimulus, which is consistent with a strong normalization in the first stage of the model. Similar results were observed in additional experiments in another animal. This surprising result suggests that nonlinearities observed in the data may be mostly implemented before the superficial layers of V1 where the VSDI signals are measured. This pair of normalization strengths was used to obtain the simulation results shown in Figure 6.

Figure 7B plots the predicted size tuning curves at the centers of the two model stages for stimuli at 5% and 100% contrasts,

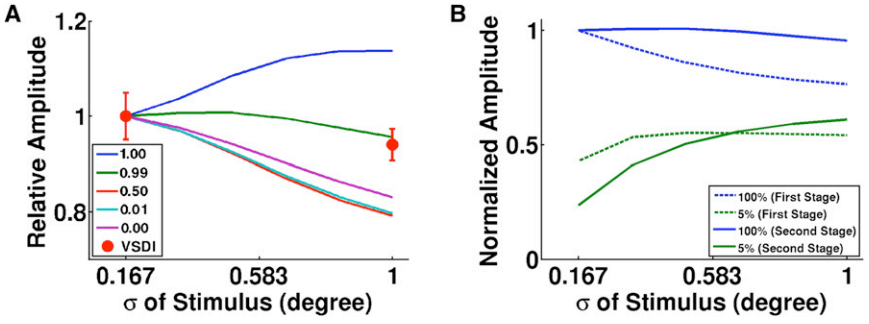


Figure 7. Size Tuning of the Model and the VSDI Responses

(A) Predictions of the normalized size tuning curves of five example combinations of normalization strengths in the two stages of the model. Each combination is labeled by the fraction of normalization in the first stage, $b_1/(b_1+b_2)$. The contrast of the stimulus was 100% and the response was measured at the center. The red dots plot the normalized responses obtained in the VSDI experiment for two stimulus sizes. The reduction of VSDI response to the large stimulus is small, which is consistent with a strong normalization in the first stage of the model.

(B) Size tuning curves at the center of the two model stages for two different stimulus contrasts (5% and 100%), with the same parameters used in Figure 6 that correspond to $b_1/(b_1+b_2) = 0.99$ in (A). The solid blue curve is the same as the green curve in (A).

using the same parameters for Figure 6. Because normalization in the second stage is relatively weak and because the receptive fields in the second stage are larger, the tuning curves for both contrasts peak at larger sizes than those in the first stage. In each stage, the peak of the tuning curve for low contrast occurs at a larger size than that for high contrast, consistent with the observations in single units in the LGN (Bonin et al., 2005) and V1 (Sceniak et al., 1999) and with previous models of normalization at the level of single neurons (Sceniak et al., 2001; Cavanaugh et al., 2002; Bonin et al., 2005).

DISCUSSION

VSDI in fixating monkeys was used to characterize the spatiotemporal dynamics of V1 population responses evoked by a small, briefly presented, visual stimulus. The VSDI signals are particularly informative because they capture responses over the entire active region in V1. The population responses exhibited systematic and unexpected nonlinear properties. At different locations, they started to rise approximately all at once, with the response at the center of the active region rising at a faster rate than those that were further away. Stimulus contrast also affected the response latencies and slopes. In contrast to the relatively complex dynamics at stimulus onset, the responses following stimulus offset fell together at approximately the same time and rate, regardless of stimulus contrast and spatial location. We also found that the spatial profile of the peak response was constant and independent of contrast.

The rich spatiotemporal dynamics observed in the responses place strong constraints on models of V1. Models that rely solely on a single-unit nonlinearity or slow lateral propagation in V1 are inconsistent with the observed properties of the VSDI responses. Instead, we find that a simple canonical normalization-based PGC model can qualitatively account for such dynamics.

We also used the PGC model to examine the degree to which nonlinearities in V1 responses are inherited from its inputs. Contrast gain-control has been used to explain many nonlinear properties of single unit responses in the retina and LGN that are also observed in V1, such as phase advance of response at high contrast (Shapley and Victor, 1978; Victor, 1987), contrast saturation (Bonin et al., 2005), and size tuning (Bonin et al., 2005). It is therefore possible that gain-control before V1 contributes significantly to the response nonlinearities in V1. On the other hand, there is some evidence that the P-cells, which provide about 80% of the input to V1, are fairly linear (Derrington and Lennie, 1984; but see Levitt et al., 2001). The PGC model predicts how the responses to a large stimulus depend on the nonlinearity in V1 and its input. Results from an additional VSDI experiment that varied stimulus size suggest that most of the gain-control for localized stimuli is implemented before the superficial layers of V1 (i.e., in the retina, LGN, and/or layer 4 in V1).

It is perhaps not surprising that contrast gain-control is implemented at various stages along the visual pathway, given its crucial role in preserving tuning characteristics of V1 neurons (except contrast tuning), while allowing high sensitivity to contrast. Potential advantages of implementing a large component of the contrast gain-control before V1 is that it could then help preserve tuning in the retina and LGN (as well as in cortex), and it could be implemented with a spatial pooling that might involve relatively fewer connections than required in V1.

Population gain-control is a simple and effective mechanism that can maintain the sensitivity and tuning of neurons; hence it is quite possible that it operates in most, if not all, sensory cortical areas. If so, then the population dynamics reported here in V1 may be observed in many other cortical areas, and the corresponding pathways might be simulated by a cascade of PGC stages.

Relationship between the Responses of a Single Neuron and a Neural Population

When the responses of a large population of neurons are pooled, the result can behave quite differently from the individual neurons that contribute to it. This idea is illustrated nicely in the size tuning behavior in our model. When all of the normalization occurs in the first stage of the model, then the units in this stage have strong size tuning. However, when these units are pooled linearly to produce the response of a unit in the second stage, this unit has much weaker size tuning (Figure 7A). The reason is that as the size of the stimulus increases, some of the units in the first stage that provide input to this unit decrease their responses due to surround suppression, while others increase their responses, because the stimulus now enters the center of their receptive field. The net effect of increasing the stimulus size is therefore much weaker in the second-stage unit than in the individual units in the first stage that provide input to it (Figure 7B). This is one factor that may explain why the VSDI responses (which measure the summed activity of a large population of V1 neurons) are only slightly lower for a large stimulus than for a small stimulus, even though strong size tuning has been observed for single neurons in V1 (Sceniak et al., 1999, 2001; Cavanaugh et al., 2002; Levitt and Lund, 2002).

A second factor is the heterogeneity in the tuning properties of the neurons within the population. For example, the size tuning for different neurons varies greatly. There is also a large range of suppression: some neurons are suppressed to spontaneous firing rate as stimulus size increased, while some neurons are not suppressed at all. Overall, more than half of the neurons are suppressed by less than 40% of their peak responses (Cavanaugh et al., 2002). When the responses from these units are pooled together, the combined tuning curve will in general be shallower than the individual curves in the population.

The above discussion illustrates how unexpected properties can emerge at the level of neural population responses (for a more general discussion see Seidemann et al., 2009). In general, in many cases it will be difficult or even impossible to predict the population responses based on a small sample of single-unit measurements. Our results on size tuning demonstrate this difficulty and emphasize the importance of direct measurements of population responses.

Possible Implementation of Divisive Population Gain-Control

A central idea of our model is that the gain is controlled through division. A key question is therefore: How is the division achieved in a neuron? It is possible that division is implemented by

a combination of different biophysical mechanisms (Kayser et al., 2001; Carandini, 2004) at different scales. At the level of individual neurons, local nonlinearities such as synaptic depression (Abbott et al., 1997; Tsodyks and Markram, 1997) have a divisive effect on the presynaptic activity, but these mechanisms are unlikely to account for the long-range effects that we observe. Connections with inhibitory interneurons could deliver the normalization signals at the population level.

Noise reduction in the membrane potential could also contribute to gain-control by reducing the likelihood of crossing spike threshold (Finn et al., 2007); however, noise reduction may itself be the result of some form of gain-control (in many systems, lowering gain lowers noise). Further investigations will be required to understand the relationship between contrast gain-control and membrane-potential noise.

Another key question is: Where do the signals that control the gain come from? In the feedforward implementation, which is illustrated in Figure 5A, the gain of the individual neuron is computed at the same level as its *input*, and is provided to the neuron at the same time as (or before) the excitation. Alternatively, in a feedback implementation that has been previously proposed (Heeger, 1992; Carandini et al., 1997), the gain is computed from the *output* of the neuron and its neighbors. In this case the gain computation can occur either at the same level, or potentially even in a subsequent stage that then sends fast feedback. Although a feedforward circuit appears to be the simplest and most parsimonious implementation of the gain-control, a mechanism that involves very rapid feedback, potentially through a specialized subset of the neurons with fast dynamics, cannot be ruled out. Additional experiments are needed to address this important question.

The PGC model assumes that conductance changes instantaneously with the input. Although this is not plausible, there is evidence suggesting such a change occurs within milliseconds (Albrecht et al., 2002). In addition, simulations of a modified model where the conductance changed with a time constant of 10 ms showed that there was no qualitative difference in the responses. The basic instantaneous model thus provides a reasonable approximation to a more realistic model.

Traveling Wave of Neural Responses

Consistent with previous VSDI studies (Grinvald et al., 1994; Jancke et al., 2004; Benucci et al., 2007), our results show that if the latency of the response is estimated from the time to half peak (t_{50}) or from the response phase (using spectral analysis), then a traveling wave would appear to originate from the center and propagate toward the periphery at a moderate speed of ~0.4 mm/ms (see Figures 2E and 3E). Although the accepted hypothesis for such spatiotemporal dynamics is that they reflect propagation of responses through slow lateral connections in V1, our results suggest that lateral connections are not necessarily the major cause for such dynamics. In fact, the lateral connections hypothesis predicts a spatial increase of response onset latency (t_{10}) that is not observed in our data. Our results are therefore inconsistent with a major contribution of slow lateral propagation to the observed dynamics in V1 in responses to small stimuli (see Supplementary Materials 2 for a quantitative analysis of lateral spread). Instead, our measurements suggest that these dynamics are the result of changes in the slope of the rising response, which could be explained as a gain-control effect.

Importantly, we show that a feedforward PGC model, in which the responses reach all locations in V1 at the same time, can account for the spatial changes in t_{50} or phase of VSDI responses shown here. In Supplementary Materials 5 we also show that the model can account for the previously reported results of Benucci et al. (2007). Note that although the PGC model can explain these results, it is currently specific to the stimuli used in our experiments; altering the stimulus properties can potentially change the dynamics of the responses. Thus, it remains to be seen if our simple feedforward model can predict the response dynamics for a wider range of stimulus conditions.

CONCLUSION

To understand the processing of arbitrary visual stimuli in the cortex, it is important to characterize the properties of V1 population responses and evaluate models that can account for them. As an initial step, we used VSDI in fixating monkeys to fully characterize the spatiotemporal dynamics of the population responses in the superficial layers of V1 evoked by a small, briefly presented, visual stimulus. The population responses exhibited several systematic and unexpected nonlinear properties that are not obvious from single unit results. We also showed that models with static nonlinearities and models with slow lateral propagation of responses in V1 are inconsistent with the observed properties of the VSDI responses. Instead, a simple canonical population gain-control model was found to qualitatively account for such dynamics. The consistency of our data with population gain-control and the advantages of such a mechanism for simultaneously providing tuning invariance and high sensitivity to weak signals suggests that population gain-control is likely to operate in most, if not all, sensory cortical areas.

EXPERIMENTAL PROCEDURES

The results reported here are based on methods that have been described in detail previously (Seidemann et al., 2002; Chen et al., 2006, 2008). Here we focus on details that are of specific relevance to the current study. All procedures have been approved by the University of Texas Institutional Animal Care and Use Committee and conform to NIH standards.

Behavioral Task and Visual Stimulus

A monkey was trained to maintain fixation while a small oriented stationary Gabor stimulus was presented on a uniform gray background. Each trial began when the monkey fixated on a small spot of light (0.1° × 0.1°) on a video display. Following an initial fixation, the Gabor stimulus was presented for 200 ms at 2.2° eccentricity, with σ of 0.167° and spatial frequency of 2.5 cycles per degree. Throughout the trial, the monkey was required to maintain gaze within a small window (<2° full width) around the fixation point in order to obtain a reward. Early fixation breaks invalidated the trials, which were not included in the analysis. Each block of trials contained 8 to 12 different contrasts from 0% (blank) to 100% presented pseudorandomly, and ten valid trials were run for each condition.

In a separate set of experiments, the σ of the Gabor stimulus was either 0.167° or 1° in each trial. The contrast of the stimulus was always 100%, and it was presented for 100 ms. The other parameters of the stimulus were the same as the experiment described above.

Analysis of Imaging Data

Imaging data were collected at 100 Hz at a resolution of 512 × 512 pixels. The size of each pixel was 37 × 37 μm^2. Our basic analysis is divided into four steps: (i) normalize the responses at each pixel by the average fluorescence at that pixel across all trials and frames, (ii) remove from each pixel a linear trend estimated on the basis of the response in the 100 ms interval before stimulus onset for each trial, (iii) remove trials with aberrant VSDI responses (generally less than 1% of the trials) (see Chen et al., 2008), and (iv) subtract the response to the blank condition from the stimulus-present conditions.

After the basic analysis described above, the spatial properties of the responses in individual trials were determined. First, the center of the spatial response of each experiment was estimated by fitting a 2D Gaussian to the average response taken over a time window of 160–260 ms after stimulus onset (shaded region in Figure 2A), for all stimulus contrasts (25% to 100%). This center was then held fixed while the average response over the same time window was fitted with a 2D Gaussian to determine the lengths of the major and minor axes and the orientation of the major axis for each trial of the experiment.

To include more trials at each contrast level in the analysis, we combined responses of five experiments from one monkey. Due to the slight difference in the setup of each experiment, the spatial responses could be translated and rotated with respect to each other. The center and average orientation of the 2D Gaussian fit of each experiment were used to transform the data so that the spatial responses aligned and overlapped in all the experiments. Data from individual experiments are similar to the combined data but noisier.

Model Definition

Processing in Each Model Unit

In the linear step of the model, each unit in a stage computes the weighted sum of the input $I(x,t)$ by cross-correlation with a Gaussian spatial receptive field:

$$A(x,t) = I(x,t) \otimes G(x),$$

where $G(x) = (1/\sigma_G\sqrt{2\pi})\exp(-0.5(x/\sigma_G)^2)$, and $\otimes$ denotes cross-correlation evaluated at x. Note that if the input is a Gaussian in space, then the weighted sum across the population will also be a Gaussian. The summation then passes through an RC circuit to produce the response. The response of the unit at x can be described by the following RC circuit equation (see Figure 5B):

$$C\frac{dV(x,t)}{dt} = A(x,t) - g(x,t)V(x,t),$$

where C is the fixed capacitance, $A(x,t)$ is the receptive field summation activity, and $g(x,t)$ is the conductance of the resistor for the stage. The conductance at each unit increases with the normalization pool activity$B(x,t)$ and is defined as

$$g(x,t) = g_0(1 + B(x,t)),$$

where g_0 is the fixed baseline conductance. For each unit the normalization activity is given by: $B(x,t) = b \cdot I(x,t) \otimes H(x)$, where b is a scaling factor that represents the strength of normalization and $H(x)$ is the Gaussian weighting function defining the normalization pool (see Figure 5B). All the parameters are the same for the units in the same stage, but they can differ between stages.

The detailed dynamics of the model are analyzed in Supplementary Materials 6.

Response Transformation

Because the response in each model stage represents average membrane potential, responses in the first stage are converted into spikes by a power function before being sent to the second stage. In other words, the input $I_2(x,t)$ that the second stage receives from the first stage is:

$$I_2(x,t) = V_1(x,t)^n,$$

where $V_1(x,t)$ is the response in the first stage. The same function is also applied to the activity in the input layer before feeding into the first stage. (Note that applying a power exponent to a Gaussian profile changes the width, but leaves the shape Gaussian.)

Simulation of the Model

The values of the parameters for the two stages in the model were estimated by fitting the responses in the model V1 to the VSDI responses. To reduce the number of free parameters, we assumed $g_0 = 1$ for both stages because it is effectively a scaling factor of the response and the capacitance. The constant delay in the input layer was chosen to be 20 ms, which was a few milliseconds shorter than the shortest latency seen in the data. The exponent n of the power function that converts membrane potential into spikes was chosen to be 2, which is similar to what we found experimentally (C.R. Palmer, Y. Chen., and E. Seidemann, 2008, Soc. Neurosci., abstract) and provides a good fit to the data. The same exponent is used in the power function in the input layer. Based on the literature suggesting that the widths of the center and surround in the afferents of V1 are about half of those in V1 (Sceniak et al., 2006), we also assumed $\sigma_{G,2} = 2\sigma_{G,1}$ and $\sigma_{H,2} = 2\sigma_{H,1}$. By assuming the width of the VSDI spatial profile to be the result of cascaded receptive field summations and the power functions, we estimated the value of $\sigma_{G,1}$. Using the difference in the rising edge slopes at different locations, we estimated the normalization pool size in the second stage, $\sigma_{H,2}$, by the procedure discussed in Supplementary Materials 4.

The remaining free parameters that needed to be estimated were C_1, C_2, b_1, and b_2. We first fitted the center's normalized contrast response function to the data by minimizing the sum of the squared error. This step enabled b_1 and b_2 to be determined separately from C_1 and C_2, because the capacitances do not affect the steady-state response in the model. After that, the normalization strengths, b_1 and b_2, were held fixed, while the capacitances were estimated by fitting the slopes of the rising and falling edges at different locations and stimulus contrasts simultaneously. The obtained parameters were $\sigma_{G,1} = 0.983$ mm, $\sigma_{H,1} = 1.386$ mm, $C_1 = 3.19$, $C_2 = 2.30$, $b_1 = 1521$, and $b_2 = 2$. The model was simulated for a 20 mm long strip (extending the black rectangular region in Figure 1C) using the Matlab function *ode45()*.

SUPPLEMENTAL DATA

Supplemental Data include modeling details, 5 figures, and references and can be found with this article online at http://www.cell.com/neuron/supplemental/S0896-6273(09)00848-4.

ACKNOWLEDGMENTS

We thank W. Bosking, C. Michelson, C. Palmer, and Z. Yang for assistance with experiments and for discussions, and T. Cakic for technical support. This work was supported by National Eye Institute Grants EY-016454 and EY-016752 (to E.S.) and EY-02688 (to W.S.G.) and by a Sloan Foundation Fellowship (to E.S.).

Accepted: August 10, 2009
Published: December 23, 2009

REFERENCES

Abbott, L.F., Varela, J.A., Sen, K., and Nelson, S.B. (1997). Synaptic depression and cortical gain control. Science *275*, 221–224.

Albrecht, D.G. (1995). Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. Vis. Neurosci. *12*, 1191–1210.

Albrecht, D.G., and Geisler, W.S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. Vis. Neurosci. *7*, 531–546.

Albrecht, D.G., Geisler, W.S., Frazor, R.A., and Crane, A.M. (2002). Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function. J. Neurophysiol. *88*, 888–913.

Albrecht, D.G., and Hamilton, D.B. (1982). Striate cortex of monkey and cat: contrast response function. J. Neurophysiol. *48*, 217–237.

Anderson, J.S., Lampl, I., Gillespie, D., and Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. Science *290*, 1968–1971.

Angelucci, A., Levitt, J.B., Walton, E.J., Hupe, J.M., Bullier, J., and Lund, J.S. (2002). Circuits for local and global signal integration in primary visual cortex. J. Neurosci. *22*, 8633–8646.

Bair, W., Cavanaugh, J.R., Smith, M.A., and Movshon, J.A. (2002). The timing of response onset and offset in macaque visual neurons. J. Neurosci. *22*, 3189–3205.

Benucci, A., Frazor, R.A., and Carandini, M. (2007). Standing waves and traveling waves distinguish two circuits in visual cortex. Neuron *55*, 103–117.

Blasdel, G., and Campbell, D. (2001). Functional retinotopy of monkey visual cortex. J. Neurosci. *21*, 8286–8301.

Bonin, V., Mante, V., and Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. J. Neurosci. *25*, 10844–10856.

Carandini, M. (2004). Receptive fields and suppressive fields in the early visual system. In The Cognitive Neurosciences III, M.S. Gazzaniga, ed. (Cambridge, MA: MIT Press), pp. 313–326.

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

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

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

Chen, Y., Geisler, W.S., and Seidemann, E. (2006). Optimal decoding of correlated neural population responses in the primate visual cortex. Nat. Neurosci. *9*, 1412–1420.

Chen, Y., Geisler, W.S., and Seidemann, E. (2008). Optimal temporal decoding of neural population responses in a reaction-time visual detection task. J. Neurophysiol. *99*, 1366–1379.

Dean, A.F., and Tolhurst, D.J. (1986). Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat striate cortex. Exp. Brain Res. *62*, 143–151.

Derrington, A.M., and Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. J. Physiol. *357*, 219–240.

Dow, B.M., Snyder, A.Z., Vautin, R.G., and Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. Exp. Brain Res. *44*, 213–228.

Finn, I.M., Priebe, N.J., and Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. Neuron *54*, 137–152.

Gilbert, C.D., and Wiesel, T.N. (1979). Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. Nature *280*, 120–125.

Gonzalez-Burgos, G., Barrionuevo, G., and Lewis, D.A. (2000). Horizontal synaptic connections in monkey prefrontal cortex: an in vitro electrophysiological study. Cereb. Cortex *10*, 82–92.

Grinvald, A., Lieke, E.E., Frostig, R.D., and Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. J. Neurosci. *14*, 2545–2568.

Grinvald, A., and Hildesheim, R. (2004). VSDI: a new era in functional imaging of cortical dynamics. Nat. Rev. Neurosci. *5*, 874–885.

Heeger, D.J. (1991). Nonlinear model of neural responses in cat visual cortex. In Computational Models of Visual Processing, M.S. Landy and J.A. Movshon, eds. (Cambridge, MA: MIT Press), pp. 119–133.

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

Hirsch, J.A., and Gilbert, C.D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. J. Neurosci. *11*, 1800–1809.

Hubel, D.H., and Wiesel, T.N. (1974). Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. J. Comp. Neurol. *158*, 295–305.

Jancke, D., Chavane, F., Naaman, S., and Grinvald, A. (2004). Imaging cortical correlates of illusion in early visual cortex. Nature *428*, 423–426.

Kaplan, E., and Benardete, E. (2001). The dynamics of primate retinal ganglion cells. Prog. Brain Res. *134*, 17–34.

Kayser, A., Priebe, N.J., and Miller, K.D. (2001). Contrast-dependent nonlinearities arise locally in a model of contrast-invariant orientation tuning. J. Neurophysiol. *85*, 2130–2149.

Levitt, J.B., and Lund, J.S. (2002). The spatial extent over which neurons in macaque striate cortex pool visual signals. Vis. Neurosci. *19*, 439–452.

Levitt, J.B., Schumer, R.A., Sherman, S.M., Spear, P.D., and Movshon, J.A. (2001). Visual response properties of neurons in the LGN of normally reared and visually deprived macaque monkeys. J. Neurophysiol. *85*, 2111–2129.

Mante, V., Bonin, V., and Carandini, M. (2008). Functional mechanisms shaping lateral geniculate responses to artificial and natural stimuli. Neuron *58*, 625–638.

Martin, K.A., and Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. J. Physiol. *353*, 463–504.

McIlwain, J.T. (1986). Point images in the visual system: new interest in an old idea. Trends Neurosci. *9*, 354–358.

Murakoshi, T., Guo, J.Z., and Ichinose, T. (1993). Electrophysiological identification of horizontal synaptic connections in rat visual cortex in vitro. Neurosci. Lett. *163*, 211–214.

Nelson, D.A., and Katz, L.C. (1995). Emergence of functional circuits in ferret visual cortex visualized by optical imaging. Neuron *15*, 23–34.

Parker, A.J., and Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. *21*, 227–277.

Purushothaman, G., and Bradley, D.C. (2005). Neural population code for fine perceptual decisions in area MT. Nat. Neurosci. *8*, 99–106.

Rockland, K.S., and Lund, J.S. (1983). Intrinsic laminar lattice connections in primate visual cortex. J. Comp. Neurol. *216*, 303–318.

Sceniak, M.P., Chatterjee, S., and Callaway, E.M. (2006). Visual spatial summation in macaque geniculocortical afferents. J. Neurophysiol. *96*, 3474–3484.

Sceniak, M.P., Hawken, M.J., and Shapley, R. (2001). Visual spatial characterization of macaque V1 neurons. J. Neurophysiol. *85*, 1873–1887.

Sceniak, M.P., Ringach, D.L., Hawken, M.J., and Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. Nat. Neurosci. *2*, 733–739.

Sclar, G., and Freeman, R.D. (1982). Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. Exp. Brain Res. *46*, 457–461.

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

Seidemann, E., Arieli, A., Grinvald, A., and Slovin, H. (2002). Dynamics of depolarization and hyperpolarization in the frontal cortex and saccade goal. Science *295*, 862–865.

Seidemann, E., Chen, Y., and Geisler, W.S. (2009). Encoding and decoding with neural populations in the primate cortex. In The Cognitive Neurosciences IV, M.S. Gazzaniga, ed. (Cambridge, MA: MIT Press).

Shapley, R.M., and Victor, J.D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. J. Physiol. *285*, 275–298.

Skottun, B.C., Bradley, A., Sclar, G., Ohzawa, I., and Freeman, R.D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior. J. Neurophysiol. *57*, 773–786.

Tsodyks, M.V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc. Natl. Acad. Sci. USA *94*, 719–723.

Van Essen, D.C., Newsome, W.T., and Maunsell, J.H. (1984). The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. Vision Res. *24*, 429–448.

Victor, J.D. (1987). The dynamics of the cat retinal X cell centre. J. Physiol. *386*, 219–246.

Witkovsky, P., Schmitz, Y., Akopian, A., Krizaj, D., and Tranchina, D. (1997). Gain of rod to horizontal cell synaptic transfer: relation to glutamate release and a dihydropyridine-sensitive calcium current. J. Neurosci. *17*, 7297–7306.

Yang, Z., Heeger, D.J., and Seidemann, E. (2007). Rapid and precise retinotopic mapping of the visual cortex obtained by voltage-sensitive dye imaging in the behaving monkey. J. Neurophysiol. *98*, 1002–1014.

Neuron, volume *64*

Supplemental Data

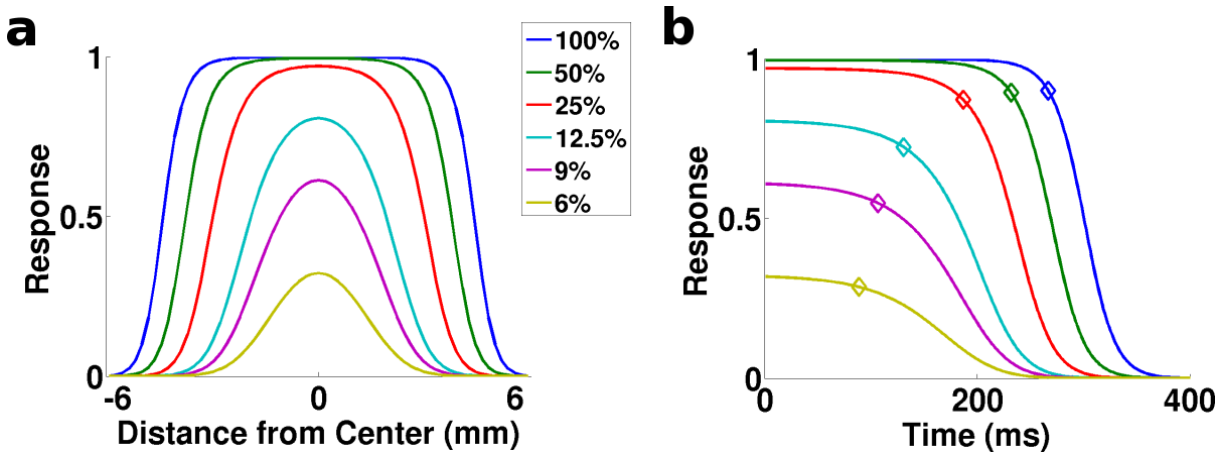
Complex Dynamics of V1 Population Responses Explained by a Simple Gain-Control Model

Yiu Fai Sit, Yuzhi Chen, Wilson S. Geisler, Risto Miikkulainen, and Eyal Seidemann

1 Static Nonlinearity Models

This section shows the predictions of a static nonlinearity model regarding the spatial profile and falling edge latency of the responses as a function of contrast. The nonlinearity was a Naka-Rushton equation, with parameters determined from the VSDI responses (Figure 1f). To obtain the spatial profile at a particular contrast, a fixed Gaussian input with σ = 2.1 mm was first scaled linearly with contrast and then passed through the Naka-Rushton equation. Supplementary Figure 1a plots the spatial profiles for the contrasts used in the VSDI experiments. As contrast increased, the region with saturated responses became larger, resulting in a wider spatial profile. To see this, note that at low contrast a model unit with its receptive field centered on the stimulus would produce a much bigger response than a peripheral unit with a receptive that only slightly overlaps the stimulus. But, at high contrasts the responses would become much closer to equal, because the response to the central unit would have already saturated at a low contrast, allowing the response of the peripheral unit to "catch up". This property is inconsistent with the VSDI responses, because the spatial profile does not depend on stimulus contrast (Figure 1e).

The simulation of the falling edges in the model was similar. A fixed sigmoidal time course using parameters obtained from the VSDI responses (Figure 4) was first scaled with contrast and then the Naka-Rushton equation was applied at each time point. Supplementary Figure 1b plots the falling edges for different contrasts. The responses at higher contrast remained saturated for a longer time, thus increasing the latency. To see this, consider the predicted response when the stimulus contrast is sufficiently high to push the output of the initial linear filter above the level that produces response saturation. In this case the response will remain saturated until the output of the linear filter drops below the point where saturation is produced. Thus, the higher the contrast the longer the time after stimulus offset until the response begins to fall below the saturated level. Such contrast dependence for the latency of falling edge of the response is not observed in the VSDI responses (see Figure 4), suggesting that the nonlinearity models cannot account for the VSDI responses.



Supplementary Figure 1: Predictions of the static nonlinearity model. (a) Spatial profile of the responses for different contrasts. The profile widens substantially as contrast increases, which is not observed in the VSDI responses (see Figure 1d). (b) Falling edge of the response for different contrast. At high contrast, the response remains saturated for a long period of time, delaying the latency (diamond symbols). In contrast, the falling edge latencies are similar for different contrasts in the VSDI responses (see Figure 2b). The static nonlinearity model is therefore inconsistent with the observed responses dynamics.

2 Modeling lateral propagation

To investigate how the latencies (t_{10}) are affected by lateral propagation, we ran simulations of a model in which the difference in the rising edges was solely due to lateral spread. We assumed that the stimulus-driven signal in each model unit had the same sigmoidal time course $A(t)$, with amplitude $c(x)$ proportional to the local contrast in the unit's receptive field. These stimulus-driven signals were propagated through lateral connections at a constant speed, which was implemented as a spatiotemporal kernel $G_l(x,t)$:

$$G_l(x,t) = \begin{cases} G(x) & \text{if } t = \dfrac{|x-c|}{v} \\ 0 & \text{otherwise} \end{cases},$$

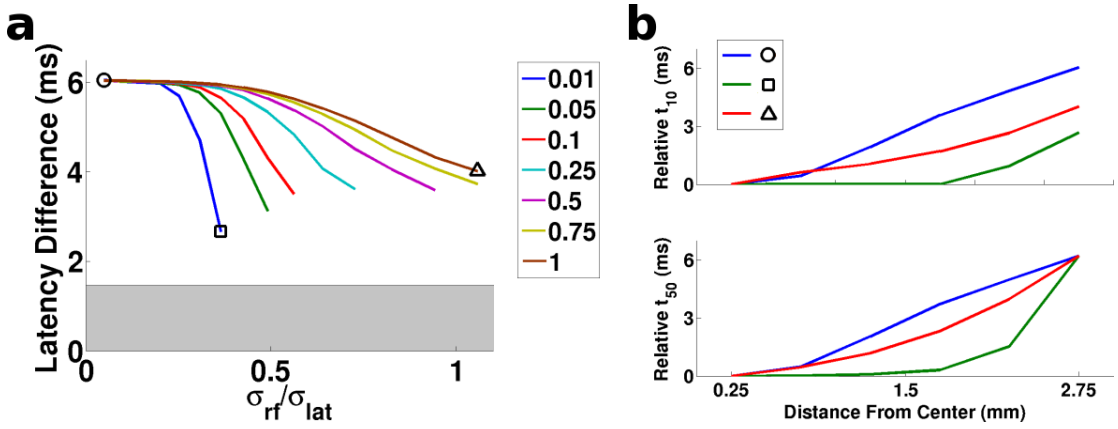
where v is the conduction speed, c is the center of the kernel, and $G(x)$ is the weighting function of lateral connections. The responses of the unit were the sum of stimulus-driven signals and the incoming lateral signals multiplied by a scaling factor w:

$$V(x,t) = c(x)A(t) + w \cdot (c(x)A(t)) \otimes G_l(x,t)$$

In such a model, the overall strength, extent, and conduction speed of the lateral connections all contribute to the time course of the observed response. In our simulations, we assumed that the spatial strength of the lateral connection $G(x)$ and the receptive field envelope followed a Gaussian profile. Given the width of the response profile (σ = 2.1 mm), the widths of the lateral connections σ_{lat} and the receptive field σ_{rf} were constrained by $\sigma^2 = \sqrt{\sigma_{stim}^2 + \sigma_{rf}^2 + \sigma_{lat}^2}$, where σ_{stim} = 0.5 mm is the width of the direct projection of the stimulus used in the experiment (stimulus width of 1/6 deg times cortical magnification factor of 3 mm/deg). For each scaling factor w and each plausible pair of values for σ_{rf} and σ_{lat}, we fitted the conduction speed v of the lateral connections so that the times to half of the peak response (t_{50}) at distances 0.25 mm and 2.75 mm differed by 6.2 ms (the experimentally observed time difference). The latencies (t_{10}) at these two locations were then determined from the fitted responses.

Supplementary Figure 2a plots the latency difference as a function of the ratio between σ_{rf} and σ_{lat} for different weights w over the range where a fit was possible. The predicted latency difference is substantial and Supplementary Figure 2b plots how the latency and time to half peak increase as a

function of distance from the center for the two extreme points with $w = 1$ in Supplementary Figure 2a (circle and triangle symbols) and the rightmost point for $w = 0.01$ (square symbol). As the ratio between σ_{rf} and σ_{lat} increased, the fitted conduction speed decreased to compensate for the wider feedforward receptive field. At the end of each curve in Supplementary Figure 2a, the speed was less than 0.01 mm/ms, which is an order of magnitude slower than the speed observed in lateral connections (0.1-0.4 mm/ms; Hirsch & Gilbert, 1991; Murakoshi, Guo, & Ichinose, 1993; Grinvald, Lieke, Frostig, & Hildesheim,1994; Nelson & Katz, 1995; Gonzalez-Burgos, Barrionuevo, & Lewis, 2000) and may not be biologically plausible. For the values of the weight w that were tested, the predicted differences were larger than the 95% confidence interval of the latency difference in the VSDI responses. This result suggests that linear summation of delayed signals from lateral connections alone is inconsistent with the dynamics of the rising edges of the VSDI responses. Changing the shape of the stimulus-driven response did not affect the result, as long as it remained sigmoid.



Supplementary Figure 2: Effects of lateral propagation on the rising edge latency. (a) The difference in latencies between the locations that were 0.25 mm and 2.75 mm away from the center as a function of the ratio between the widths of the receptive field and lateral connections, for different values of the lateral weight scaling factor w. Each curve is plotted for the range where a fit to the time to half peak was possible. Shaded region shows the 95% confidence interval (truncated at 0) between the latencies at the two locations in the VSDI responses. (b) The increase in latency (top) and time to half peak (bottom) as a function of distance from the center for three example points indicated by different symbols in (a).

References

Gonzalez-Burgos, G., Barrionuevo, G., & Lewis, D. A. (2000). Horizontal synaptic connections in monkey prefrontal cortex: an in vitro electrophysiological study. *Cerebral Cortex, 10*(1), 82-92.

Grinvald, A., Lieke, E. E., Frostig, R. D., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *Journal of Neuroscience, 14*(5), 2545-2568.

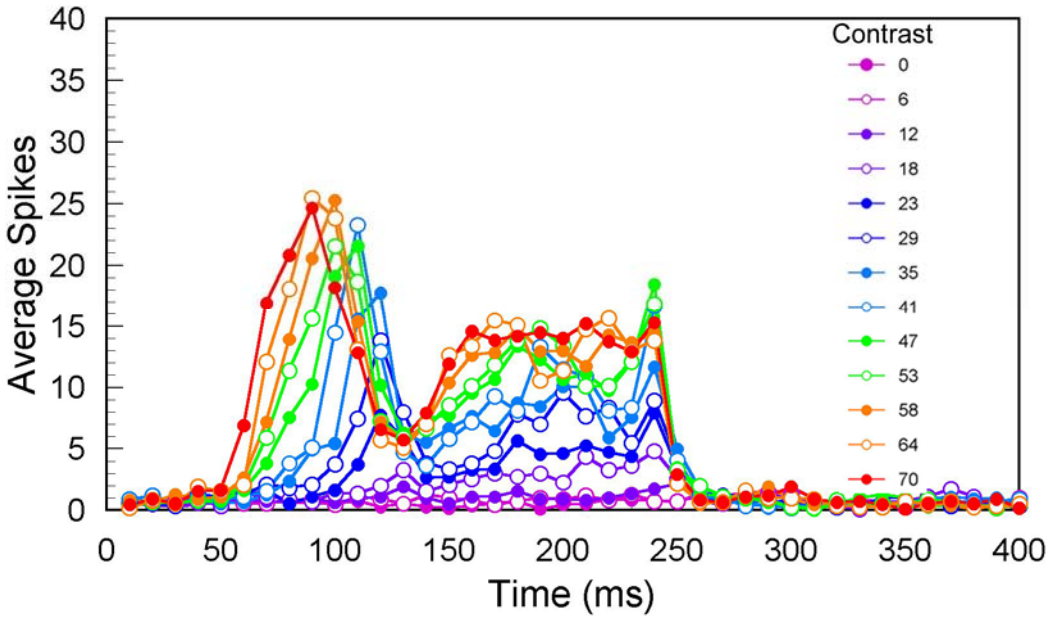
Hirsch, J. A., & Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience, 11*(6), 1800-1809.

Murakoshi, T., Guo, J. Z., & Ichinose, T. (1993). Electrophysiological identification of horizontal synaptic connections in rat visual cortex in vitro. *Neuroscience Letters, 163*(2), 211-214.

Nelson, D. A., & Katz, L. C. (1995). Emergence of functional circuits in ferret visual cortex visualized by optical imaging. *Neuron, 15*(1), 23-34.

3 Asymmetrical affect of contrast on onset and offset latencies in single neurons

We have shown here that a strong prediction of the PGC model is that onset latency decreases as the contrast of a transiently presented stimulus is increased, but that the offset latencies are largely independent of the contrast of the stimulus (see Figure 6C). This prediction also holds for similar contrast normalization models developed to explain single neuron responses, but to our knowledge has not been tested in the literature. Although a thorough test of this prediction is warranted, there is some evidence that the prediction holds approximately for single neurons in the primary visual cortex of cat. Supplementary Figure 3 shows the average response of the 4 complex cells measured in Geisler, Albrecht & Crane (2007) to sine-wave grating patches of various contrasts, presented for 200 msec. As contrast is increased the latency of the response onset decreases; however, at stimulus offset the responses decline approximately simultaneously at all contrasts. The same trend was seen for simple cells. This behavior is qualitatively consistent with the PGC model.



Supplementary Figure 3. Average post-stimulus-time histograms of 4 complex cells in the primary visual cortex of the cat, for sine-wave grating stimuli of various contrasts, presented for 200 ms. The grating stimuli were confined to the classical receptive field and drifted at 5 Hz (one cycle in 200 ms). (Adapted from Geisler et al., 2007.)

Reference

W.S., Albrecht D.G. & Crane A.M. (2007) Responses of cortical neurons to transient changes in local contrast and luminance. *Journal of Neuroscience*, 27, 5063-5067.

4 Estimating normalization pool size

The width of the normalization pool affects how the time constant of the rising edge changes across space. Given the measured rising edge time constants τ_c at the center c and τ_p at the periphery that is p mm from the center, we can estimate the size of the pool in the second stage ($\sigma_{H,2}$). This procedure gives the lower bound for $\sigma_{H,2}$, because it does not take the difference in the slopes of the input to the second stage into account; if there is already some difference in the input, $\sigma_{H,2}$ would need to be larger to offset the difference.

First, the time constant at center is given by

$$\tau_c = \frac{C}{g_0(1+B(c))}$$
$$B(c) = \frac{\tau}{\tau_c} - 1,$$

where $\tau = C/g_0$, which is the time constant of the falling edge. Since both τ and τ_c are known, the above equation can be evaluated to give a numerical value. Similarly, $B(p) = \frac{\tau}{\tau_c} - 1$ and can be computed.

Note that B is a Gaussian with width $\sqrt{\sigma_1^2 + \sigma_{H,2}^2}$, where σ_1 is the width of the spiking activity in the first stage, which is related to the width of the input, σ_x, and the width of the receptive field in the first stage, $\sigma_{G,1}$, by $\sigma_1^2 = (\sigma_x^2/2 + \sigma_{G,1}^2)/2$. $B(c)$ and $B(p)$ are the values at their corresponding points. Dividing $B(c)$ by $B(p)$, the width of the normalization pool in the second stage can be estimated:

$$\frac{B(c)}{B(p)} = \frac{1}{\exp\left(-\frac{p^2}{2(\sigma_1^2 + \sigma_{H,2}^2)}\right)}$$

$$\sigma_1^2 + \sigma_{H,2}^2 = \frac{p^2}{2\ln\left(\frac{B(c)}{B(p)}\right)}$$

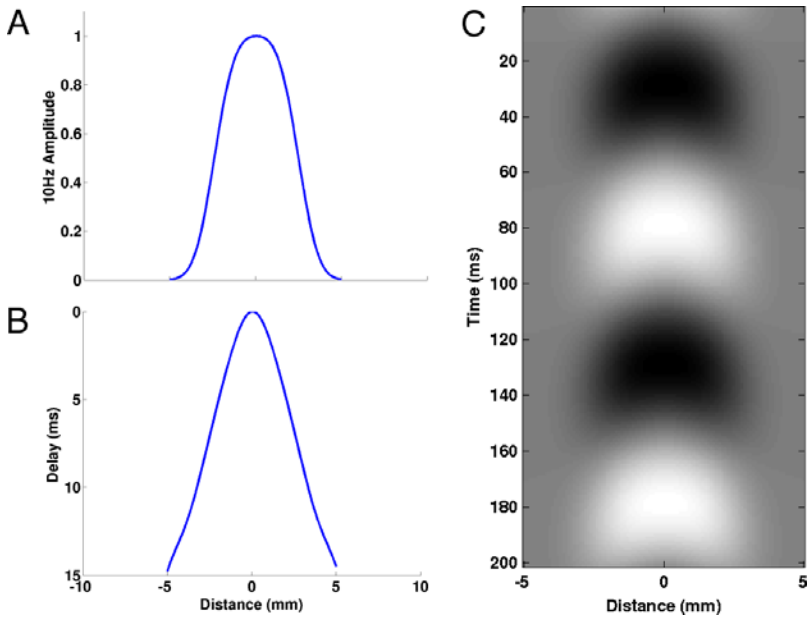
$$\sigma_{H,2} = \sqrt{\frac{p^2}{2\ln\left(\frac{B(c)}{B(p)}\right)} - \sigma_1^2}\,.$$

Note that the above computation is independent of the normalization strength b.

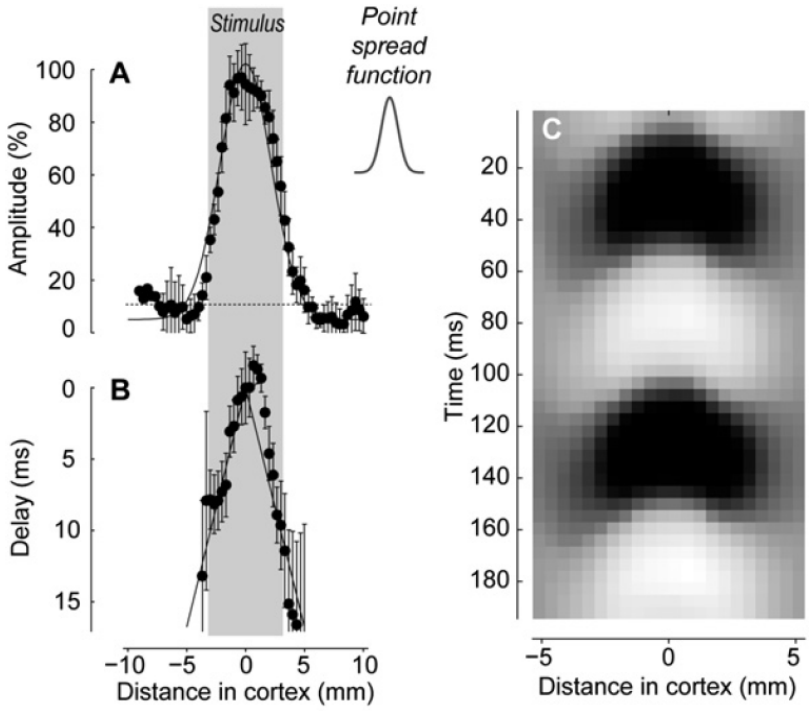
5 Simulation results for the experiments reported in Benucci, Frazor, & Carandini (2007)

We have used the PGC model to simulate the experiments described in Benucci, Frazor, & Carandini (2007), which reports traveling waves of neural responses (as measured by their phases) in the V1 of anesthetized cats.

The stimuli used in theses experiments were square gratings that reversed contrast sinusoidally. In the simulation, theses gratings were represented as rectangular envelopes with appropriate widths on the input layer of the model so that the widths of the Gaussian point spread function (1.1 mm) and spatial response (2.2 mm) match those reported in that paper. To represent the contrast reversing of the visual stimulus the amplitude of the rectangular envelope modulated sinusoidally at 5 Hz and was full-wave rectified. The model parameters used were $\sigma_{G,1} = 0.5\,\text{mm}$ (resulting in a point spread of $\sqrt{0.5^2 + (0.5 \times 2)^2} = 1.1\,\text{mm}$ in the second stage of the model), $\sigma_{H,1} = 0.9\,\text{mm}$, $C_1 = 2.46$, $C_2 = 1.53$, $b_1 = 121$, and $b_2 = 20$. Note that these values are specific for this simulation to provide a good fit to the data; they are different from those used in the main text of the paper.



Supplementary Figure 4: Simulation results of the model. (A) Spatial profile of the 10 Hz component of the responses. (B) The delay of the responses (computed from their phases) at different distances from the center. The delay at a distance of 5 mm is about 15 ms. (C) Spatiotemporal response of the model, band-pass filtered at 7-13 Hz as in Benucci et al., (2007).



Supplementary Figure 5: Results of Benucci et al., (2007). The subfigures are arranged in the same way as in Supplementary Figure 4. The predictions of the model are very similar to the data, suggesting that normalization can account for the observed traveling waves.

Supplementary Figure 4 shows the simulation results, which are similar to those in Benucci et al. (2007) (Supplementary Figure 5). The similarity of the predictions of the model with the data suggests that, with appropriate parameters, a feedforward PGC model can account for the increase in the phase of the responses across space reported in Benucci et al. (2007). However, as with the small Gabor stimuli used in our study, this increase in phase in the simulated data is due to the difference in gain across space rather than lateral propagation of responses. Note that the responses shown in Supplementary Figure 4C are band-passed filtered at 7-13 Hz to mirror the analysis performed in that paper. Such filtering distorts the shape of the response time courses, increasing the apparent symmetry between the rising and the falling edges in that figure. In addition, note that because here the stimulus contrast rises and falls-off smoothly (instead of rising and falling-off abruptly as in our study), the time courses of the simulated rising and falling edges depend on position. Specifically, at the center, the time constant (gain) is the smallest, and hence the latencies of the rising and falling edges are shortest and their slopes are steepest.

Reference

Benucci, A., Frazor, R. A. & Carandini, M. (2007). Standing waves and traveling waves distinguish two circuits in visual cortex. *Neuron, 55*, 103-117.

6 Temporal dynamics and steady state of the model

The dynamics of a model area can be illustrated by analyzing its response to a step input with a Gaussian spatial profile. For the rising edge, the input is zero initially. Then the input $I(x,t)$ becomes positive and remains constant. As a result, both the receptive field summation $A(x,t)$ and conductance $g_0(1+B(x,t))$ are constant in time (after stimulus onset) and the dynamics for the unit at x can be simplified to

$$C\frac{dV(x,t)}{dt} = A(x) - g_0(1+B(x))V(x,t),$$

which has the solution

$$V(x,t) = \frac{A(x)}{g_0(1+B(x))}\left(1-\exp\left(-\frac{g_0}{C}(1+B(x))t\right)\right).$$

This equation simply means that the rising edge of the response is the process of charging a capacitor with time constant $C/\left(g_0(1+B(x))\right)$. For a Gaussian input, $B(x)$ is also a Gaussian. Its value thus falls off gradually away from the center. Since the time constant is inversely related to $B(x)$, the time constant at the center will be shorter than those at the periphery. The responses $V(x,t)$ thus rise faster at the center.

As time progresses, responses will reach steady state $A(x)/\left(g_0(1+B(x))\right)$. To analyze how this function changes with contrast for a Gaussian input, note that for a given unit x, $A(x)$ can be written as $r(x)I(x)$, and similarly, $B(x)=s(x)I(x)$. This gives the steady state response

$$V(x,t) = r(x)\frac{I(x)}{g_0(1+s(x)I(x))},$$

which is a saturating function of $I(x)$ that approaches $r(x)/\left(g_0 s(x)\right)$ as $A(x)$ increases. The response therefore saturates when the stimulus contrast is high.

For the falling edge responses of the model, consider a step input that switches from a positive constant to zero. Both A and B will therefore become zero, resulting in the following dynamics:

$$C\frac{dV(x,t)}{dt} = -g_0 V(x,t).$$

The equation has a simple solution

$$V(x,t) = V_o(x)\, e^{-\frac{g_o}{C}(t-t_o)},$$

where $V_o(x)$ and t_o are the responses and the time at the offset of input, respectively. This solution has two important implications. First, the responses at all locations will start to decay all at once after the stimulus disappears, regardless of stimulus contrast and V_0. Second, the responses will decay with the same time constant C/g_0. In addition, note that the time constant of the rising edge $C/\left(g_0(1+B(x))\right)$ is smaller than C/g_0. This result explains why the slopes of the falling edges were predicted to be smaller than those in the rising edges, as observed in the data.

Neuron
Previews

The V1 Population Gains Normalization

Elad Ganmor,[1] Michael Okun,[1] and Ilan Lampl[1,*]
[1]Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel
*Correspondence: ilan.lampl@weizmann.ac.il
DOI 10.1016/j.neuron.2009.12.015

In this issue of *Neuron*, Busse et al. describe the population response to superimposed visual stimuli while Sit et al. examine the spatiotemporal evolution of cortical activation in response to small visual stimuli. Surprisingly, these two studies of V1 report that a single gain control model accounts for their results.

Orientation selectivity is the hallmark of the primary visual cortex (V1). When this property was discovered more than 40 years ago by Hubel and Wiesel, it was thought that the selectivity of cortical cells results only from the organization of feedforward inputs from the visual thalamus. Today we know that the response of cortical cells might be strongly affected by inputs from the entire visual field.

Hubel and Wiesel showed that in V1 cells, the response evoked by lines or bars at specific angles (orientation) is much greater than the response to circular spots of light (Hubel and Wiesel, 1962). However, the first-order cortical neurons that display this property (simple cells) receive their afferent inputs from neurons of the lateral geniculate nucleus (LGN) of the thalamus, which are not orientation selective. How then can this behavior be explained? Hubel and Wiesel suggested a simple model in which simple cells receive feedforward inputs from several LGN neurons with aligned receptive fields. When stimulated by an elongated stimulus, aligned with the collective receptive field structure of these thalamic cells, they are activated simultaneously, causing a large response in simple cells.

Although many predictions of this feedforward model were confirmed experimentally, other predictions failed. One major discrepancy is the observation that the width of orientation tuning curves in V1 is independent of the stimulus contrast (Sclar and Freeman, 1982), a phenomenon called contrast invariance. Since the firing of geniculate cells increases monotonically with contrast, the feedforward model predicts that as contrast is increased, stimuli further away from the preferred orientation will evoke sufficient depolarization to cause firing. Thus, the tuning curve of V1 neurons is expected to widen when contrast increases (Ferster and Miller, 2000).

Another experimental observation not explained by the simple feedforward model is the strong suppression of responses to a stimulus at the preferred orientation by an orthogonal stimulus, even if the orthogonal stimulus by itself evokes no response (Priebe and Ferster, 2006). The feedforward model predicts that the response to a combination of stimuli is merely the sum of the responses to each individual stimulus.

Subsequently, new models were proposed to account for the experimental findings described above. Roughly, they can be described as belonging to two categories: feedforward models extending the original model of Hubel and Wiesel, and models incorporating feedback inputs. Normalization models, also known as contrast gain control models (Albrecht and Geisler, 1991; Heeger, 1992), belong to the first category. In these models, two pathways determine the response of a cortical neuron. One is a specific filter defined by the neuron's selectivity to the stimulus, as in the feedforward model of Hubel and Wiesel. The second pathway integrates less selective inputs from a wider visual field, and serves as a normalization background. That is, the response of the cell is a result of dividing the input from the first pathway by the input via the second (gain control) pathway (Figure 1). At the level of single cells, contrast gain control models were found to be successful in explaining several key features of visual processing, in particular contrast invariance and cross-orientation suppression.

Visual information, however, is represented by the joint activity of many neurons. Population models of V1 have relied mostly on data gathered from single neurons, yet neural populations may display qualitatively different behaviors than the units that comprise them. For example, a population of contrast invariant neurons may not be contrast invariant in itself (see below). Therefore, it is not clear whether the aforementioned models, developed to describe the responses of single neurons, can be successfully applied to neuronal populations.

A study by Busse et al. (2009) in this issue of *Neuron* efficiently addresses this question, using multielectrode arrays to record from many neurons in cat V1. Busse and colleagues characterized the tuning curves of multiple simultaneously recorded neurons, and examined how their responses to a superposition of two oriented stimuli sum together. The population response was defined as the average firing rate of neurons grouped according to their preferred orientation. Interestingly, the authors found that a simple normalization model can account for their results.

Initially Busse and her colleagues verified that the population response is contrast invariant and therefore a simple normalization model, composed of a product of a tuning curve and a contrast gain function, may be applied to the population response. How can a population response not exhibit contrast invariance when single neurons are known to be contrast invariant? Consider, for example, a population in which sharply tuned neurons have high contrast thresholds whereas widely tuned cells have low thresholds. In such a population, as contrast is increased, more sharply tuned neurons are recruited, resulting in a sharpening of the population tuning curve. However, the authors find that contrast sensitivity and tuning width are

independent of each other in the population, giving rise to contrast invariant orientation tuning at the population level.

To investigate V1 population responses to more complex stimuli, cats were presented with a sum of two oriented gratings (a plaid), where the contrast of each grating was varied separately. The authors found that the population responses to a combination of stimuli can range from equal weight summation of the responses to the individual stimuli, to a winner-take-all regime in which only one stimulus is represented while the other is virtually ignored. The factor that determines how the responses are summed is the contrast of the respective stimuli. For similar contrast values, an equal weight summation takes place, whereas for large differences in contrast only the response to the high contrast grating is retained. What model can account for this wide range of weight combinations?

Busse, Wade, and Carandini demonstrate that a normalization model for single neurons can be adapted to describe the population response. In this model responses are nonlinearly scaled by their contrast, summed, and then normalized (divided) by the overall contrast of both stimuli. Dividing by the overall constant results in suppression among concurrent stimuli. The nonlinear scaling with contrast results in equal weight summation in the case of similar contrast but amplifies the difference when dissimilar contrasts are used, leading to winner-take-all competition. Thus, the model captures cross-stimulation suppression, and the smooth transition between equal weight summation and winner-take-all, without requiring a change in assigned weights for different stimuli.

Is the normalization preformed by V1 cells, or is it already present in the subthreshold input to these neurons? Busse and colleagues found that the normalization model provides a good fit to the average local field potential (LFP) responses of the entire population to plaid stimuli, suggesting that population subthreshold activity in V1 neurons can be described by the same normalization model.

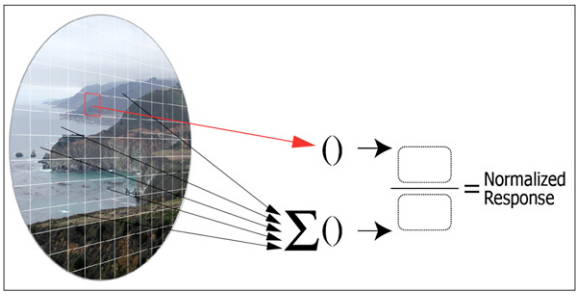


Figure 1. The Normalization Model for Visual Processing
The response of a cortical neuron to visual stimulation is determined by two pathways: (1) excitatory input from its classic receptive field (red), and (2) a gain control component, modulated by a wider range of visual inputs (black). The overall response is determined by the quotient of the two components (each component is also subject to some nonlinearities, not shown here for simplicity).

A second study in this issue of *Neuron*, by Sit et al. (2009), supports Busse and colleagues' finding that population normalization operates already at the subthreshold potential range of upper cortical layers. The authors of this study investigated the responses of V1 neurons to small oriented stimuli in awake monkeys using voltage sensitive dye imaging (VSDI), which reflects changes in membrane potential (Grinvald and Hildesheim, 2004), and examined their spatiotemporal evolution. In spite of the vast difference in methods used in the two studies, Sit and colleagues report that a closely related model accounts for their experimental results.

Previous studies in primate V1 showed cortical activation far beyond the retinotopic mapping of the stimulus (Grinvald et al., 1994). This wide spatial activation is commonly attributed to a spread of activity via lateral connections among cortical neurons. The present study suggests that this may not be the case. Sit and colleagues show that the latency of subthreshold responses of V1 cells, as measured by the VSDI signal, is constant regardless of the distance from the retinotopic center of activation. This result does not agree with the model of lateral propagation, because this model predicts that the latency should increase with distance due to synaptic delays.

A natural candidate to account for the constant latency is the classic feedforward model. If the activation observed using VSDI is due to feedforward connections, then clearly we would expect no difference in latency. However, further results dismiss the feedforward model. Specifically, the authors find that the area and spatial profile of cortical activation are invariant to contrast. The classic feedforward model predicts that an activation spatial profile will grow wider as contrast is increased, and thus cannot explain the data.

Finally, Sit et al. explore a population gain control model. In their two-stage model, each neuron receives feedforward excitation from neurons in its receptive field pool. In addition, its conductance is modulated by neurons in a normalization pool that in particular contains the receptive field of the neuron. Increased activity in the normalization pool results in higher conductance in its target neurons and therefore has two major effects on their response: the amplitude is decreased and the time constant is reduced, leading to faster dynamics. The feedforward connectivity explains the constant latency of responses across the entire active region, while the normalization, or gain control, accounts for the invariant profile of spatial spread when contrast varies. The increase in conductance accounts for another experimental observation—the slope of activation increases with proximity to the activation center. In the model, higher conductance for units positioned near the response center, due to higher activity in their normalization pool, reduces their time constant and increases the slope.

Importantly, because the spiking latency depends on the slope of activation, the finding that subthreshold response latency, captured by the VSDI, is independent of the distance from the center of activation region might not be observed via spike measurements. Hence, the use of VSDI reveals an important property of cortical response that proved essential for the conclusions of Sit et al.

Although both studies provide compelling support for contrast gain control in visual processing, it is not immediately clear how the models presented in the two studies are related. Hence, it is worth noting that the conductance model on which the two-layer circuit of Sit et al. is

based was proposed as a possible biophysically plausible implementation of a normalization operation (Carandini and Heeger, 1994), such as the one in Busse et al.

What biophysical mechanisms may be responsible for divisive gain control? Different studies have addressed this question. One candidate mechanism is short-term synaptic depression (Freeman et al., 2002). Widely tuned, visually evoked cortical shunting inhibition may also contribute to contrast normalization. However, intracellular recording studies in vivo of inhibitory tuning curve profiles and changes in evoked conductance in response to plaid stimuli (Priebe and Ferster, 2008) found no support for this view.

Divisive gain control might support higher-level aspects of visual processing beyond the responses of V1 neurons to relatively simple stimuli. Some studies debate the role of normalization in redundancy reduction and efficient coding (Schwartz and Simoncelli, 2001; Shi et al., 2006), while others suggest that changes in visual processing (sensitivity, gain, etc.) induced by shifts in attention may be explained by a modulation of the input signal by an attentional filter followed by normalization (Reynolds and Heeger, 2009).

Clearly, the functional implications of contrast gain control for downstream visual areas and the contribution of different biophysical mechanisms to its implementation are still open questions. Hopefully, further research and analysis of how large populations process complex stimuli may shed light on these issues.

REFERENCES

Albrecht, D.G., and Geisler, W.S. (1991). Vis. Neurosci. *7*, 531–546.

Busse, L., Wade, A.R., and Carandini, M. (2009). Neuron *64*, this issue, 931–942.

Carandini, M., and Heeger, D.J. (1994). Science *264*, 1333–1336.

Ferster, D., and Miller, K.D. (2000). Annu. Rev. Neurosci. *23*, 441–471.

Freeman, T.C.B., Durand, S., Kiper, D.C., and Carandini, M. (2002). Neuron *35*, 759–771.

Grinvald, A., and Hildesheim, R. (2004). Nat. Rev. Neurosci. *5*, 874–885.

Grinvald, A., Lieke, E.E., Frostig, R.D., and Hildesheim, R. (1994). J. Neurosci. *14*, 2545–2568.

Heeger, D.J. (1992). Vis. Neurosci. *9*, 181–197.

Hubel, D.H., and Wiesel, T.N. (1962). J. Physiol. *160*, 106–154.

Priebe, N.J., and Ferster, D. (2006). Nat. Neurosci. *9*, 552–561.

Priebe, N.J., and Ferster, D. (2008). Neuron *57*, 482–497.

Reynolds, J.H., and Heeger, D.J. (2009). Neuron *61*, 168–185.

Schwartz, O., and Simoncelli, E.P. (2001). Nat. Neurosci. *4*, 819–825.

Sclar, G., and Freeman, R.D. (1982). Exp. Brain Res. *46*, 457–461.

Shi, J., Wielaard, J., and Sajda, P. (2006). Conf. Proc. IEEE Eng. Med. Biol. Soc. *1*, 4991–4994.

Sit, Y.F., Chen, Y., Geisler, W.S., Miikkulainen, R., and Seidmann, E. (2009). Neuron *64*, this issue, 943–956.

Discreet Charm of the GABAergic Bourgeoisie: Superconnected Cells Conduct Developmental Symphonies

Marianne Case[1,*] and Ivan Soltesz[1,*]
[1]Department of Anatomy and Neurobiology, University of California, Irvine, Irvine, CA 92697, USA
*Correspondence: marianne.case@uci.edu (M.C.), isoltesz@hs.uci.edu (I.S.)
DOI 10.1016/j.neuron.2009.12.013

In an exciting study in the December 4th issue of *Science*, Bonifazi and colleagues demonstrated the existence and importance of exceedingly rare but unusually richly connected cells in the developing hippocampus. Manipulating the activity of single GABAergic hub cells modulated network activity patterns, demonstrating their importance for coordinating synchronous activity.

Much to the chagrin of our latte-drinking, sushi-eating, Volvo-driving liberal friends all over, networks in the real world are decidedly not egalitarian but rather aristocratic in nature. Indeed, the disproportionate influence of rare superconnected hubs is well-known in technological, biological, and social networks, including aviation grids (such as LAX and JFK), biochemical reaction pathways (such as pyruvate and ATP), and the proverbial old boys' networks. For neuroscience in particular, hub-like connectors are considered to be of great potential significance because networks with such aristocratic flavor have been predicted by theoretical studies to represent a clever compromise between fast computation, economy of wiring, and robustness against random deletions (Buzsáki et al., 2004; Bullmore and Sporns, 2009). However, while we have thoroughly defined neuronal networks lacking superconnected neurons (such as that of