Behavioral/Systems/Cognitive

Benefits of Contrast Normalization Demonstrated in Neurons and Model Cells

Kate S. Gaudry and Pamela Reinagel

Division of Biology, Neurobiology Section, University of California, San Diego, La Jolla, California 92093-0357

The large dynamic range of natural stimuli poses a challenge for neural coding: how is a neuron to encode large differences at high contrast while remaining sensitive to small differences at low contrast? Many sensory neurons exhibit contrast normalization: gain depends on the range of stimuli presented, such that firing-rate modulation is not proportional to contrast. However, coding depends strongly on the precision of spike timing and the reliability of spike number, neither of which can be predicted from neural gain. The presumption that contrast normalization is associated with maintained coding efficiency remained untested. We report that, as contrast decreases, responses are more variable and encode less information, as expected. Nevertheless, these changes can be small, and information transmission is even better preserved across contrasts than rate modulation. The extent of contrast normalization is correlated with the extent to which information transmission is preserved across contrasts. Specifically, normalization is associated with maintaining the bits of information per spike rather than bits per second. Finally, we show that a nonadapting model can exhibit both contrast normalization and the associated information preservation.

Key words: adaptation; contrast gain control; information theory; coding; LGN; thalamus

Introduction

Given that a neuron can only produce a limited range of firing rates, the sensitivity of its firing rate to stimulus strength will determine the range of stimuli it can effectively discriminate. There is a broad range of contrasts present in natural scenes (Ruderman and Bialek, 1994; van Hateren, 1997) (supplemental Fig. 1, available at www.jneurosci.org as supplemental material), suggesting that it would be advantageous to adapt sensitivity to the prevailing contrast (Atick and Redlich, 1992; Schwartz and Simoncelli, 2001). Contrast normalization has long been known to occur in the early visual system of vertebrates (Shapley and Victor, 1978, 1981; Shapley and Enroth-Cugell, 1984; Benardete et al., 1992; Shou et al., 1996; Benardete and Kaplan, 1997; Smirnakis et al., 1997; Brown and Masland, 2001; Chander and Chichilnisky, 2001; Kim and Rieke, 2001, 2003; Kremers et al., 2001; Baccus and Meister, 2002; Solomon et al., 2004; Jin et al., 2005). Analogously, adaptation to the variance of higher-order stimulus features is found in visual cortex (Ohzawa et al., 1985; Bonds, 1991; Heeger, 1992; Kohn and Movshon, 2003) and many other sensory systems. For example, in the fly visual system, the firing-rate sensitivity of H1 neurons to motion scales with motion contrast (Brenner et al., 2000; Fairhall et al., 2001; Heitwerth

et al., 2005). Model simulations verify that this scaling optimizes the information available at the level of firing rate (Brenner et al., 2000).

The relationship between contrast normalization and neural coding remains unclear, however, because information is not coded only by firing rate. In the lateral geniculate nucleus (LGN), for example, both the reliability and temporal precision of spikes are essential for the encoding of high-contrast white-noise visual stimuli (Reich et al., 1997; Reinagel and Reid, 2000; Liu et al., 2001; Reinagel and Reid, 2002). As a result of lower signal-to-noise ratio, one might expect responses to be less temporally precise, less reliable from trial to trial, and less informative as contrast decreases. It is difficult, however, to predict the magnitude and significance of these effects. Contrast normalization measures firing rate properties and thus does not itself reveal the effect of contrast on neural coding, which depends on properties such as reliability and precision in addition to firing rate.

It is also not known whether contrast normalization requires an active adaptation process. Response properties of neurons change gradually after a change in contrast, and the speed of these changes has attracted theoretical interest (DeWeese and Zador, 1998; Fairhall et al., 2001). The biophysical properties of neurons also change with stimulus contrast (Carandini and Ferster, 1997; Sanchez-Vives et al., 2000; Kim and Rieke, 2001; Baccus and Meister, 2002). Conversely, there is experimental evidence that spike-generation mechanisms play a role in contrast normalization (Kim and Rieke, 2001, 2003; Zaghloul et al., 2005). Recent theoretical results show that nonlinear systems can exhibit contrast normalization without any active changes; fixed nonlinear properties of the cell (such as the threshold and saturation) can be sufficient to produce contrast normalization (Borst et al., 2005; Yu and Lee, 2005; Yu et al., 2005). In the models shown previ-

ously to exhibit contrast normalization, optimization of information transfer across contrasts was either not found (Yu and Lee, 2005) or not tested. These studies used firing-rate models that do not consider the reliability and precision of real neurons, and information transmission was not analyzed at the level of spike timing.

Materials and Methods

Surgical preparation. Cats were anesthetized initially with ketamine HCl (20 mg/kg, i.m.), followed by sodium pentothal (2–4 $\text{mg} \cdot \text{kg}^{-1} \cdot \text{h}^{-1}$, i.v., supplemented as needed). Animals were ventilated using an endotracheal tube. Electrocardiogram, electroencephalogram, temperature, expired CO_2, and oxygen in blood were continually monitored. All surgical and experimental procedures were in accordance with National Institutes of Health and United States Department of Agriculture guidelines and were approved by the University of California, San Diego Institutional Animal Care and Use Committee.

Electrical recording. We report results from 41 LGN relay cells recorded from the A laminas of the LGN of anesthetized cats. Parylene-coated tungsten electrodes (AM Systems, Everett, WA) were inserted through a 0.5-cm-diameter craniotomy over the LGN. Recordings were amplified, filtered, and digitized at 10 kHz sampling rate (CED micro 1401 and Spike2, version 5.12a; Cambridge Electronic Design, Cambridge, UK). Waveforms were analyzed off-line to isolate single-unit responses (Fee et al., 1996).

Visual stimulation. Stimuli were spatially uniform and presented on a custom-built light-emitting diode array. To create matched stimuli at all contrasts, we began with a random binary stimulus of 125 frames/s for 10 s. The same binary sequence was scaled about the mean to obtain three contrast conditions (11, 33, and 100%, in which contrast is defined as the SD of the luminance over the mean.) Stimuli of the three contrasts were interleaved and presented between 10 and 128 times each. We analyzed only the last 5 s of the response to each 10 s stimulus. The mean luminance was the same for all contrasts and was well within the photopic range (55 cd/m^2). These stimuli produced the fast but not the slow form of adaptation for most cells in our population (data not shown) (cf. Baccus and Meister, 2002). We also presented cells with a 1 s, binary stimulus at 10 different contrasts, repeated 40 times at each contrast (used in Fig. 2e,f).

Comparisons between cell types. To classify cell types, spatiotemporal receptive fields were mapped using reverse correlations of responses to a white-noise (m-sequence) stimulus. Most cells were also classified as X or Y based on observations of null phases or frequency doubling of contrast reversing stimuli modulated at 1, 2, or 4 Hz. We used both contrast-reversing sine gratings and contrast-reversing bipartite field stimuli. The latter consisted of two hemifields sinusoidally modulated in counterphase; the vertical division between the hemifields was varied, presenting 10 cycles for each location of the division. The population of 41 cells consisted of 18 ON cells (5 X ON, 9 Y ON, and 4 ON cells unclassified with respect to X/Y), 21 OFF cells (10 X OFF, 7 Y OFF, and 4 OFF cells unclassified with respect to X/Y), and two Y cells that exhibited ambiguous (possibly ON/OFF) responses. The latter two cells could not be fit with our linear–nonlinear model and were not included in analyses involving gain, but the cells were included for other analyses such as reliability, precision, and information rate.

Estimation of neural gain with a linear–nonlinear cascade. For each cell at each contrast, we began by estimating the filter as the spike-triggered average. The filters for the three contrasts were normalized by the amplitude of their first peaks. The stimulus was convolved with the corresponding filter to create a generator potential, $g(t)$. We compared this generator potential with the observed probability of spiking at each time bin. For each cell, we fit the observed input–output function at 100% contrast to the following sigmoid equation:

$$y = \exp(A - e^{-Gx+S}). \qquad (1)$$

In Equation 1, the variables A, G, and S describe the amplitude, slope, and horizontal offset of the nonlinear function, respectively. We then fit the input–output functions at the other two contrasts holding the amplitude

(A) constant for all contrast conditions for each cell. We excluded a dataset from our analysis if there were <100 spikes observed in the response, but even cells with low firing rates could usually be included by simply recording more trials to obtain sufficient numbers of spikes. We also excluded a dataset from additional analysis if it could not be well fit by a sigmoid (if the R^2 value associated with the sigmoid fit was <0.90). For each cell at each contrast, we then define the gain as G from Equation 1. Our results did not depend critically on how the filters were scaled; similar results were found by normalizing by the peak-to-peak amplitude of the spike-triggered average.

This method of estimating gain has the advantage that it allows for differing shapes of the nonlinearity, which we observed in our data. In our population of cells, the horizontal offset (S) increased as contrast increased for some cells, decreased for others, and remained constant for still others. The offset change was not correlated with information preservation (data not shown).

Nevertheless, to assess whether our method of gain estimation importantly influenced our conclusions, we replicated our entire analysis using an alternative framework for estimation of the gain, namely, fixing the nonlinear function at all contrasts and scaling the filters as needed to match the empirical relation between $g(t)$ and the observed probability of spiking (Chichilnisky, 2001). The two approaches are conceptually similar but subject to different practical limitations. Our results were qualitatively the same in both analyses (data not shown).

Spike-count variability measure. The trial-by-trial variability in spike count was measured by the Allan Factor (AF):

$$\text{AF} = \frac{< N_i - N_{i-1})^2 >}{2\mu}, \qquad (2)$$

where N_i is the spike count in trial i, μ is the mean spike count over all trials, and $< \ldots >$ denotes the average result across all consecutive trials at one point in time in the repeated flickering stimulus. The Allan Factor was computed separately for all nonoverlapping 5 ms windows of the stimulus, and we report the average over these windows. In windows with no spikes in any trial ($\mu = 0$), AF is undefined.

The Allan Factor is closely related to the more commonly used Fano Factor, but instead of calculating the variance as the average squared difference in spike count with respect to the average spike count across all trials, the numerator is the average difference in spike count with respect to the spike count during the preceding trial. The Allan Factor is preferred because it is less sensitive than the Fano Factor to slow changes in neural activity over the course of long recordings. All trends reported remained the same, however, if the Fano Factor was used instead of the Allan Factor (data not shown).

Responses to repeated stimuli were divided into nonoverlapping 5 ms bins. [Trends were the same using 10, 50, 100, 250, 500, and 1000 ms bins (data not shown).] The Allan Factor was computed from the distribution of spike counts across trials separately in each time bin, and these values were averaged across time bins for each cell within each contrast condition.

We performed an alternative analysis using firing events (Berry et al., 1997), such that, instead of calculating the Allan Factor from the number of spikes within each bin, the calculation was performed on the number of spikes within each peristimulus time histogram (PSTH) peak. We either averaged the Allan Factors across firing events and compared the average across contrasts, or compared the Allan Factors between contrasts on a peak-by-peak basis (data not shown). Both yielded similar results to the time-binned method.

It is known that variability is low when instantaneous firing rate is high because of the regularizing effects of refractoriness (Kara et al., 2000). To test whether the observed increase in mean firing rate (see Fig. 2e) or in rate modulation (see Fig. 2f) was sufficient to explain the observed decrease in variability at high contrast (see Fig. 4a), we also compared the Allan Factors for a subset of time bins from each contrast that fell within a narrow range of firing rates. This range was chosen separately for the high–medium and high–low contrast comparisons, because often there was no one firing rate sufficiently represented at all three contrasts. We required a minimum $n = 10$ samples at each contrast. We used 5 ms bins,

such that, at most, one PSTH peak was contained within a bin. Because the mean rates could not be perfectly matched between two contrasts, we erred in the direction of including higher firing rate bins for the lower contrast, so that our bias if any would be to attribute less variability to low-contrast responses. Even when firing rate was matched, the Allan Factor of the cells increased as contrast decreased ($p < 0.0001$). We obtained similar results when comparing the variability of individual firing events (PSTH peaks) that contained the same average number of spikes (data not shown).

A minority of cells had super-Poisson variability (Allan Factor > 1), which was attributable to bursting (Kara et al., 2000). We use interspike interval criteria to identify these bursts, which we attribute to low-threshold calcium channels (Lu et al., 1992). We define the burst probability as the ratio of the burst frequency over the response frequency, in which a response is either a burst or a single spike. Across all of our conditions (41 cells × 3 contrasts), 33 of 123 had burst probability >5%; this subset included all 20 conditions for which AF > 1.05 (supra-Poisson variability).

Temporal precision measures. For each cell at each contrast condition, we measured the trial-to-trial jitter in spike trains by the width of the peak of the average cross-correlation between sequential trials. The cross-correlation reveals the probability of a spike in one trial at times immediately surrounding the time of a spike in the preceding trial. Because the timing of spikes is relatively reproducible in these data (temporal jitters on the order of milliseconds) and the time between spikes is relatively long (tens to hundreds of milliseconds), we observe sharp peaks centered at $t = 0$ in the cross-correlation between trials. We defined precision as the half-width of this peak at half-maximum. Thus, if spike times had a normal distribution of σ about the mean time, our measure would assign a precision of $(-2 \times \ln(0.5) \times \sigma^2)^{1/2}$. For analysis of temporal jitter, we included only those peaks that were sufficiently sampled (enough spikes) to accurately estimate the peak width. Specifically, our criteria were that the cross-correlation curve was sufficiently smooth to allow unambiguous identification of the peak, and this peak was at $t = 0$. To determine smoothness in an automated way, the cross-correlogram was smoothed with a five-point moving average, and the smoothness is defined as 1 minus the summed squared error between the smoothed and raw correlograms. We used a criterion of smoothness ≥ 0.75, which eliminated from this analysis all correlograms that were too noisy to accurately estimate the peak.

Because all measures of temporal precision have advantages and limitations, we repeated our entire analysis using several different measures, with similar results. For example, we also determined the average width of identified firing events (PSTH peaks) within each contrast and compared this measure of precision across contrasts. In another analysis, we compared the peak widths between contrasts on a peak-by-peak basis, which relied on our ability to identify corresponding PSTH peaks in the responses to the same stimulus time course presented at different contrasts. Both yielded similar results to the cross-correlation method (data not shown). We note that all of these definitions of precision are sensitive to the duration of the firing event as well as the jitter in the time of onset. Thus, our estimate is an upper bound on the precision of the timing of the first spike in each discrete firing event.

Contrast normalization index. The greater the change in contrast, the larger change in gain (sensitivity) would be required to compensate. Thus, we express the magnitude of gain change relative to the magnitude of contrast change and call this the contrast normalization index κ:

$$\kappa = \frac{G_{\text{lower}}/G_{\text{higher}} - 1}{C_{\text{higher}}/C_{\text{lower}} - 1}, \tag{3}$$

where G is neural gain, and C is stimulus contrast. This index has the property that $\kappa = 0$ if there is no change in gain, and $\kappa = 1$ if the gain decreased by the same factor that the contrast increased.

Mutual information measures. We calculated the visual information in spike trains by a direct entropy method (Strong et al., 1998), implemented exactly as by Reinagel and Reid (2000). Briefly, we represented LGN responses as time-binned spike trains. In our analysis, we varied $\delta\tau$ from 1 to 16 ms; results for $\delta\tau = 2$ ms are shown. The value in a time bin

was set to zero if no spikes occurred during that time interval or one if a single spike occurred during that time interval. Because of the refractory period of the cells, the occurrence of two spikes in the same 2 ms time bin was so rare as to be negligible.

We analyzed the information in words (short strings of bins) and varied the number of bins in the words, L. For each word length, we measured two forms of word entropy: the average noise entropy, $<H_{\text{noise}}>$, which reflects the trial-to-trial variability of words when the stimulus was fixed, and the average total entropy, $<H_{\text{total}}>$, which reflects the variability of words across all stimuli in the ensemble. The mutual information between the visual stimulus and the spike train is defined as $I = <H_{\text{total}}> - <H_{\text{noise}}>$. Finally, we define the coding efficiency of the cell as the mutual information divided by the total entropy (I/H_{total}).

In detail, $H_{\text{noise}}(t)$ was calculated from the distribution of words at a fixed time t relative to stimulus onset across all repeated trials of the same sample of the stimulus. We performed a separate calculation of $H_{\text{noise}}(t)$ for many different values of t (separated by one bin). We then averaged over t to get the average noise entropy $<H_{\text{noise}}>$. We performed equally many separate calculations of $H_{\text{total}}(i)$, but the set of words was instead selected using a different time t from each trial. Twice the number of words were used for each single estimate of $H_{\text{total}}(i)$ to compensate for the approximately twofold difference in entropy. We averaged over i to get the average total entropy $<H_{\text{total}}>$. We computed information as a function of word length. Results are shown for a word length of 1 bin (2 ms), but all trends were the same at all word lengths tested.

For each entropy estimate, we determined how our estimate of H converged as we used increasing fractions of the data and then corrected for finite data size according to the method of Strong et al. (1998). We fit a second-order polynomial to 1/(fraction of data) versus the entropy estimate. We used this polynomial to extrapolate to infinite data only if the resulting correction was <10% and the second-order term <1%. If data did not meet these criteria, a linear fit was used to correct for finite data size (maximum 15% correction), and we indicate our reduced confidence in these data by open symbols in Figures 5 and 6. All of our results were qualitatively unchanged if we did not perform correction for finite data size (data not shown). Once the information rate was estimated in this way, we separately analyzed the results in units of raw information rate (bits per second), information per spike (bits per spike, by dividing the information rate by the firing rate), or coding efficiency, defined as the fraction of the entropy of the cell that carries visual information ($I/<H_{\text{total}}>$).

Nonadapting LGN model. Model cells were implemented as described by Keat et al. (2001) (Fig. 1, schematic). Briefly, the model first convolves the stimulus with a linear filter. The generator potential is equal to the convolved stimulus plus noise; the amplitude and time constant of this noise are defined by the model parameters σ_a and τ_A, respectively. A spike is generated whenever the generator signal crosses a threshold, θ. Each time a spike occurs, a negative afterpotential is added to the generator potential, such that the threshold is crossed repeatedly during sustained excitatory stimuli. The amplitude, time constant, and variability in the amplitude of the negative afterpotential are defined by the model parameters B, τ_P, and σ_b, respectively. We generated a set of model cells using all possible combinations of the following parameters: $\theta = 0.1$; $B = 3, 5,$ or 7; $\tau_P = 20, 35,$ or 50; $\tau_A = 20$; $\sigma_a = 0.01, 0.16, 0.31,$ or 0.61, $\sigma_b = 0.02, 0.15,$ or 0.28; $F = F_X$ or F_Y. The filter functions, F_X and F_Y, were taken from the X ON and Y OFF cat LGN data of Keat et al. (2001), their Figure 8c.

We note that this model was chosen because it was more successful at replicating the dependence of neural responses on contrast than other models we tested. In particular, we found that a linear–nonlinear Poisson refractory model also exhibited contrast normalization and changes in spike-count reliability (data not shown), but that model failed to replicate changes in temporal precision and latency that are captured by the model we use here. A more detailed discussion of the modeling results will be presented elsewhere (Gaudry and Reinagel, 2007).
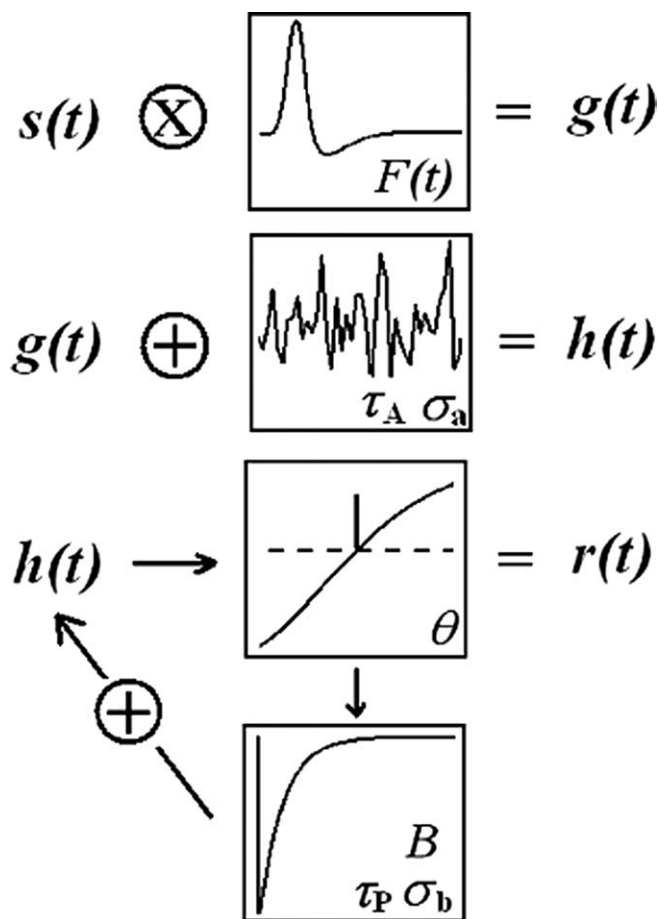
Figure 1. Schematic of generative model used to predict spike trains from stimuli. We simulated neural responses with a model originally described by Keat et al. (2001). First row, The stimulus, $s(t)$, is convolved with a filter, $F(t)$, to produce $g(t)$. Second row, A noise signal is added to $g(t)$ to produce the generator potential, $h(t)$. The parameters τ_A and σ_a determine the time constant and amplitude of this noise term. Bottom rows, When the generator potential crosses a threshold, θ, a spike occurs. Each time a spike occurs, a negative afterpotential is added to the subsequent generator potential. The parameters B, τ_P, and σ_b determine the amplitude, the time constant, and variability in the amplitude of the negative afterpotential.

Results

We recorded responses of LGN neurons to full-field flickering binary white-noise stimuli with each of three contrasts: 100, 33, or 11% (see Materials and Methods). We determined that this range is representative of temporal contrasts found in natural stimuli (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). The responses of one representative cell are shown in Figure 2. Consistent with previous reports (Reinagel and Reid, 2000; Liu et al., 2001), the high-contrast white-noise stimulus elicited sparse responses with discrete firing events separated by periods with no spiking activity (Fig. 2a). When the same stimulus sequence was scaled to lower contrasts, the firing events appear to be weaker and noisier, as expected (Fig. 2b,c). Nevertheless, even at 11% contrast, there are discrete and precise peaks in the time-varying firing rate (Fig. 2d). For this cell, both firing rate (Fig. 2e) and rate modulation (Fig. 2f) increase with contrast. Nevertheless, the rate modulation of the LGN neurons does not increase linearly, consistent with the predictions of contrast gain control (Shapley and Victor, 1978; Benardete et al., 1992). The term "contrast gain control" refers to a specific phenomenon in the vertebrate retina that encompasses not only gain changes (contrast normalization) but also changes in temporal

dynamics and spatial summation, which are not explored here. In this work, we restrict our analysis and conclusions to one aspect of contrast gain control, namely contrast normalization in the temporal dimension. If we can identify principles relating contrast normalization (i.e., normalization to input variance) to general properties of neurons, such principles could be applicable to a wide variety of sensory neurons, including those at later stages of processing, in other species, and in other sensory modalities.

Contrast normalization in LGN neurons

We measured the gain of LGN neurons by fitting the data to a linear–nonlinear cascade, to test whether cells matched their gain to the stimulus contrast. Contrast normalization in the retina has been extensively studied in this framework (Hunter and Korenberg, 1986; Chander and Chichilnisky, 2001; Chichilnisky, 2001; Kim and Rieke, 2001; Rieke, 2001; Baccus and Meister, 2002). In the analysis shown here, we estimated a linear filter $F(t)$ for each cell in each contrast condition by the normalized spike-triggered average stimulus (see Materials and Methods). Filters obtained from one Y OFF cell during high-contrast stimuli (red) and medium-contrast stimuli (blue) are shown in Figure 3a. We found a range of absolute latencies and filter durations, but filter durations in the range of 50 ms were typical in our data. This is somewhat faster than other reports in the literature (Benardete and Kaplan, 1999), but these data are not in conflict. The absolute latency and duration of the filter depends on details of the experimental preparation (species and temperature for in vitro retinal studies) and the stimulus (luminance, contrast, and temporal bandwidth), which differ among these studies.

By scaling the linear filters to the same amplitude at all contrasts, we can estimate the gain, G, of the neuron from the slope of a sigmoidal input–output function (nonlinearity) that is empirically measured from the data (Fig. 3b) (for details and alternative analyses, see Materials and Methods). The gain indicates the extent to which the probability of firing of the neuron will increase for a given increase in stimulus strength. Therefore, a higher gain corresponds to an increased neural sensitivity.

When the stimulus contrast was decreased, we found that the gain of the neurons increased (all symbols are above unity line in Fig. 3c), indicating that the sensitivity (slope of the input–output function) increased. Nevertheless, the gain changed by less than the change in contrast. In other words, a threefold drop in contrast was not compensated by an exactly threefold boost in gain (thin line, compare with circles and diamonds) and likewise for a ninefold change in contrast (dotted line, compare with triangles). We defined a contrast normalization index, κ, such that $\kappa = 0$ if there was no gain change and $\kappa = 1$ if the gain of the neuron increased by the same factor that the stimulus contrast decreased (see Materials and Methods). The average contrast normalization, κ, was ~ 0.5, consistent with previous reports in the retina and LGN (Shapley and Enroth-Cugell, 1984; Benardete et al., 1992; Chander and Chichilnisky, 2001; Kim and Rieke, 2001; Kremers et al., 2001; Baccus and Meister, 2002; Zaghloul et al., 2005). Some individual cells exhibited almost no gain change ($\kappa \approx 0$), whereas others came close to compensatory scaling ($\kappa \approx 1$). Even within cell type, there was considerable variability across cells.

Several studies have reported that different cell types of the early visual system differ in the degree of contrast gain control. Previous studies in the retina, for example, found stronger contrast gain control in Y cells than X cells of the cat retina (Shapley and Victor, 1978). Contrast gain control was reported to be stronger for OFF than ON cells in salamander retina but stronger
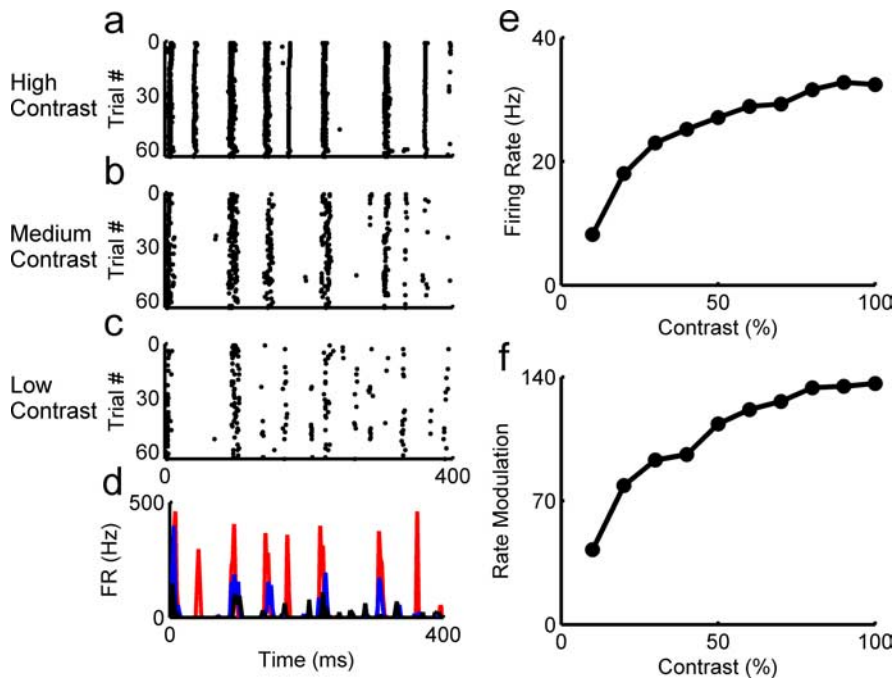
Figure 2. Dependence of responses on contrast. Responses of one Y ON cell to full-field flicker of different temporal contrasts. *a*, Raster of responses to 64 repeats of the same 100% contrast stimulus. *b*, *c*, Responses to the same stimulus sequence scaled to 33 and 11% contrast, respectively. *d*, Time-varying firing rates (FR) in 2 ms time bins from data in *a–c*. *e*, Dependence of firing rate on contrast. *f*, Effect of contrast on the modulation in firing rate (SD of the firing rate across time bins, calculated using 1 ms bins).

system could either enhance or reduce the expected change (Berry and Meister, 1998; Kara et al., 2000).

To quantify the variability in spike count, we calculated the Allan Factor of neural responses in the high- (100%), medium- (33%), and low- (11%) contrast conditions. Like the more familiar Fano Factor, the Allan Factor has a value of zero for a deterministic process, increases as variability increases, and has a value of one for a Poisson process (see Materials and Methods). Consistent with previous observations (Kara et al., 2000; Reinagel and Reid, 2000), most LGN neurons had sub-Poisson variability at high contrast (Allan Factor < 1). Responses to lower-contrast stimuli had more spike-count variability (Allan Factor increased in 80 of 82 comparisons; $p < 0.001$ by Wilcoxon's sign rank test). Nevertheless, most cells remained sub-Poisson even at 11% contrast (Fig. 4*a*). This finding, based on within-cell comparisons, is consistent with recent results based on population comparisons in the primate retina (Uzzell and Chichilnisky, 2004). We did not, however, find significant differences between ON and OFF cells with respect to spike-count variability or its dependence on contrast, as were reported by Uzzell and Chichilnisky (2004).

We measured the variability of spike timing by the temporal jitter of spikes across trials (see Materials and Methods). In our data from 100% contrast stimuli, temporal jitter was <2 ms for 61% of cells and <1 ms for 29% of cells, comparable with the precision reported previously for high-contrast white-noise stimuli (Reinagel and Reid, 2002). As the contrast decreases, the decreasing signal-to-noise ratio might be expected to produce increasing jitter in spike timing. Compared with the jitter at high contrast, nearly all cells had more temporal jitter at 33% contrast (20 of 24 comparisons; $p < 0.002$) and still more at 11% contrast ($p < 0.002$) (Fig. 4*b*). These results agree with population results from the primate retina (Uzzell and Chichilnisky, 2004) that spike time precision increases with effective contrast, but we did not find significant differences between ON and OFF cells with respect to precision or the dependence of precision on contrast, as were reported by Uzzell and Chichilnisky (2004). Note that, in our data, most cells still had temporally precise responses (<10 ms jitter) even at the lowest contrast. In a few cells (7%), submillisecond precision was retained even at 11% contrast.

for ON cells in primate retina (Chander and Chichilnisky, 2001). Both rapid and slow forms of contrast gain control were stronger in the primate magnocellular pathway than in the parvocellular pathway (Benardete et al., 1992; Benardete and Kaplan, 1999; Kremers et al., 2001; Solomon et al., 2004). In our dataset, the distributions of contrast normalization indices were not significantly different between X and Y cells for any pair of contrasts compared. We found that OFF cells had stronger contrast normalization than ON cells. In high- to medium-contrast comparisons, κ was 0.56 ± 0.25 for OFF cells versus 0.37 ± 0.20 for ON cells ($p = 0.028$); in high- to low-contrast comparisons, κ was 0.41 ± 0.20 for OFF cells versus 0.25 ± 0.12 for ON cells ($p = 0.014$); the difference was not statistically significant when comparing medium- to low-contrast ($p = 0.12$).

Variability in number and timing of spikes
Contrast normalization is thought to help maintain sensitivity under different contrast conditions. However, the gain change (which depends on the sensitivity of firing rate to stimuli) cannot predict changes in the temporal precision of responses or the trial-by-trial variability in spike number, both of which are important for determining the rate of information transmission in the LGN (Reich et al., 1997; Reinagel and Reid, 2000; Liu et al., 2001).

When a visual stimulus is presented repeatedly, there is some variability from trial to trial in the number of spikes within any time bin of the response. It is known that this spike-count variability (Fano Factor, i.e., variance divided by mean) is lower during high instantaneous firing rates, which can be explained by the regularizing effects of refractoriness (Berry and Meister, 1998; Keat et al., 2001; Uzzell and Chichilnisky, 2004; Zaghloul et al., 2005). Given that both mean rate (in some cells) and response modulation (in all cells) increased with contrast, we predicted a decrease in variability at high contrast. Nevertheless, this was not guaranteed to be the case; changes in the refractoriness of the

Contrast normalization preserves information
In most cells, responses to lower contrast had a lower firing rate and an increased variability in both spike timing and spike number. Thus, one might expect a decrease in the information rate, as measured by the mutual information between a stimulus and a neural response (Berry et al., 1997; Reich et al., 1997; Reinagel and Reid, 2000; Liu et al., 2001; Freed, 2005). Depending on the detailed structure of the neural code, information transmission could be more or less robust to contrast changes.

Therefore, we measured the mutual information I for each cell at all three contrasts (Strong et al., 1998; Reinagel and Reid, 2000) (see Materials and Methods). The information rate was significantly lower at lower contrasts, whether measured in bits per
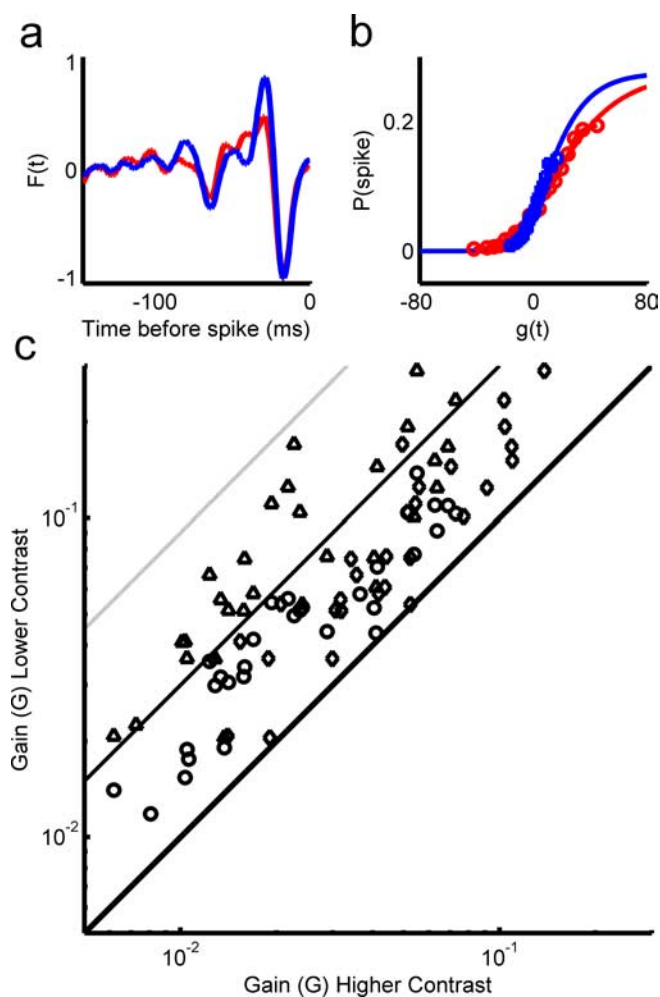
Figure 3. Gain changes with contrast. ***a***, Linear filters fit to high- (red curve) and medium- (blue curve) contrast responses from one Y OFF cell. The filters were normalized by the amplitude of their first peak. ***b***, Observed probability of spiking versus the generator potential, $g(t)$, for the same cell as in ***a***. The curves show the best-fit sigmoids (Eq. 1). The gain at the high contrast was 0.63 times that at medium contrast. ***c***, Summary of gain changes in all cells. Each symbol compares the gain of one cell between 100 and 33% contrast ($\bigcirc$), 33 and 11% contrast ($\diamond$), or 100 and 11% contrast ($\triangle$); the gain at the higher contrast is shown on the horizontal axis. Constant gain (no gain control) is indicated by the thick, solid line ($x = y$). The thin, solid line indicates a threefold gain change ($\bigcirc$, $\diamond$). Dashed line indicates ninefold gain change ($\triangle$).

second (supplemental Fig. 2, available at www.jneurosci.org as supplemental material) or in bits per spike (symbols are below the solid diagonal line in Fig. 5a), yet the information rate did not decrease by the same factor as the contrast (symbols are above their respective dashed lines). The mutual information remained substantial at 11% contrast for some cells (as high as 30.8 bit/s or 3.1 bits/spike). The extent to which information was preserved at low contrast varied considerably in our population, with no correlation to cell type.

It has long been known that contrast normalization mitigates the effect of contrast at the level of firing rate. Specifically, the modulation of the firing rate over time does not decrease by the same factor that contrast decreases. We asked whether the conservation of information rate was comparable, on a cell-by-cell basis, with the conservation of firing-rate modulation. To compare the information transmission of the same cell at two contrasts, we used the ratio of the information in bits per spike at the lower contrast condition divided by that at higher contrast condition ("information ratio"). Similarly, to compare the firing rate

modulation (see Fig. 2f), we define a "rate modulation ratio" between the two contrasts. First, we find that the extent of information conservation is highly correlated with the extent of modulation conservation, as one might expect. Second, we find information is conserved at least as well as modulation is. Indeed, we find that information coded per spike was slightly but significantly more contrast independent than was the modulation of firing rate (most symbols are above the diagonal line in Fig. 5b) ($p < 0.05$ for filled symbols). (For additional statistics of correlations among information rate, coding efficiency, firing rate, spike-count variability, and temporal jitter, see Materials and Methods.)

It is widely assumed that contrast normalization serves to maintain the fidelity of neural coding across contrasts. This conclusion does not necessarily follow because information transmission depends strongly on the temporal precision and trial-by-trial reliability of responses, which also change with contrast. We exploited the diversity among cells to test whether neurons with stronger contrast normalization were better able to preserve information transmission across contrasts. We find that the magnitude of contrast normalization, κ, is significantly correlated with the conservation of information across contrasts in terms of bits per spike (Fig. 5c) ($R^2 = 0.43$ for filled symbols; $p < 0.001$). This correlation ($R^2 = 0.32$; $p < 0.001$) was also found when information was calculated using longer words of 8 bins (16 ms).

Normalization is specifically associated with contrast-independent coding efficiency (in terms of entropy or spikes) as opposed to contrast-independent coding rate (in terms of absolute information rate). Above we reported a correlation between normalization and the conservation of information rates in units of bits per spike. We interpret bits per spike as an estimate of the energy cost of information transmission. The extent of contrast normalization was comparably well correlated with the conservation of coding efficiency (bits of mutual information per bits of entropy; $R^2 = 0.29$; analysis not shown). The raw rate of information transmission (bits per second) was also relatively contrast independent in some cells (see supplemental data, available at www.jneurosci.org as supplemental material), but the degree of conservation of information rate in bits per second was not correlated with the degree of normalization ($R^2 = 0.05$; $p > 0.05$; data not shown).

The primary results of this study are stated above and shown in Figure 5c. For completeness, we report here several other trends and correlations in our data. The fractional coding efficiency ($I/<H_{total}>$) declined significantly with contrast but remained as high as 0.39 at low contrast for some cells (supplemental Fig. 2c, available at www.jneurosci.org as supplemental material). The coding efficiency of some cells declined only modestly with contrast (efficiency ratios of up to 0.93), whereas others declined dramatically (ratios as low as 0.18). The efficiency ratio was higher on average than the information ratio, indicating that efficiency was even more invariant to contrast. The information rate in bits per second was highly correlated with the firing rate modulation ($R^2 = 0.62$; $p < 0.0001$). The spike-count variability (Fig. 4a) was inversely related to firing rate and therefore negatively correlated with information rate ($R^2 = 0.37$; $p < 0.001$). The temporal jitter of spike times (Fig. 4b) was not significantly correlated with information rate per second ($R^2 = 0.05$; $p = 0.07$) but was significantly negatively correlated with both the information per spike ($R^2 = 0.26$; $p < 0.0001$) and the coding efficiency ($R^2 = 0.26$; $p < 0.0001$).
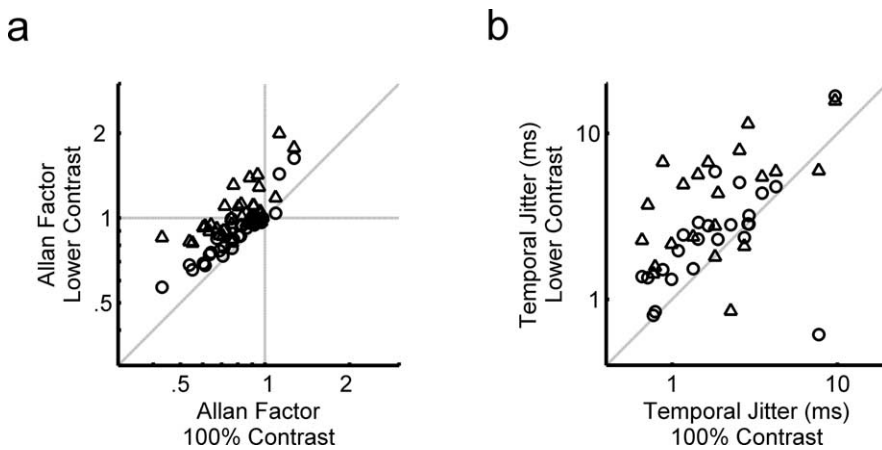
Figure 4. Variability decreases with increasing contrast. *a*, Spike-count variability (Allan Factor). Each symbol indicates the result for the responses of one LGN neuron to 100% contrast (*x*-axis) versus either 33% (○) or 11% (△) contrast (*y*-axis). *b*, Spike-timing jitter (see Materials and Methods). Symbols as in *a*.
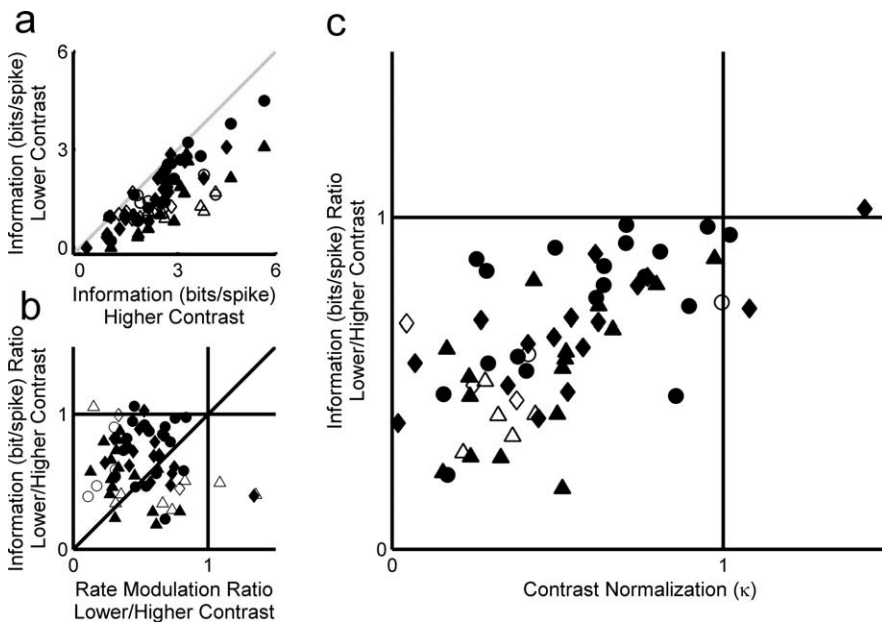

Figure 5. Visual information in LGN responses. *a*, Each symbol indicates the mutual information, in units of bits per spike, of one LGN cell at two contrasts. Results are compared between 100 and 33% contrast (○, ●), 100 and 11% contrast (△, ▲), and 33 and 11% contrast (◆, ◇). Filled symbols indicate cases in which the correction for finite data size was small and linear for both contrasts (see Materials and Methods). Information rates were significantly higher at higher contrast ($p < 0.01$ for both 100 to 33% and 100 to 11%). *b*, The ratio of information rates (in bits per spike) is plotted against the ratio of rate modulation (as defined in Fig. 2*f*). Symbols as in *a*. *c*, Each symbol indicates the information ratio for a given cell versus the contrast normalization index, κ, in the same data (symbols as in *a*).

Normalization in a nonadapting model

The results presented above are consistent with the hypothesis that cells actively adapt their sensitivity to optimize information transfer, but nonlinear systems with fixed parameters can also exhibit contrast normalization (Borst et al., 2005; Yu and Lee, 2005; Yu et al., 2005). We implemented a nonlinear model (see Fig. 1) that has been shown to successfully reproduce responses of LGN neurons to white-noise stimuli (Keat et al., 2001). We created a large population of model cells by varying the model parameters (see Materials and Methods). We then presented the model with the same stimuli presented to LGN neurons above. The parameters for each model cell remained fixed across all contrasts to determine whether parameter changes were required for contrast normalization to occur.

We find that, for this nonadapting model, the gain, *G*, increased as the contrast decreased (Fig. 6*a*). Depending on parameters, some model cells showed little or no contrast normalization (κ ≈ 0), whereas the gain of other model cells increased by nearly the same factor that the contrast decreased (κ ≈ 1). Thus, the range of contrast normalization we found in the LGN can also be produced by a fixed nonlinear system.

Contrast normalization arises in the model in part from a compression caused by spiking threshold. In addition, refractoriness acts to saturate the firing rate during the excitatory periods of the high-contrast stimuli, which effectively decreases the gain at high contrast. Noise opposes this influence to linearize neural responses. As noise increases, the instantaneous firing rate is less explained by the stimulus fluctuations (gain is reduced), and this effect is more pronounced at low contrasts.

For the present purposes, it is not important how the parameters affect the extent of contrast normalization. To test whether the normalization found in this nonadapting nonlinear model is associated with a functional benefit at the level of information transmission, it is sufficient to find parameter settings that produce a range of normalization indices comparable with that found in LGN neurons and test for covariation of this with the information ratio. A detailed exploration of how contrast normalization arises in this model will be presented elsewhere (Gaudry and Reinagel, 2007).

Like LGN cells, the information rate of the model cells also decreased when the contrast decreased (Fig. 6*b*). Again, the magnitude of this decline in information depended on the model parameters, such that some model cells lost nearly all information at low contrasts, whereas other model cells transmitted nearly the same information across contrasts. Across our population of model cells, we find that contrast normalization is correlated with the information ratio (Fig. 6*c*) ($R^2 = 0.74$; $p < 0.0001$). This correlation is remarkably similar to the one observed in LGN cells (compare blue, black, and green symbols with gray symbols). Therefore, contrast normalization without active adaptation is a viable candidate mechanism for how LGN neurons preserve information transmission across contrasts.

Discussion
Contrast dependence of neural reliability and precision
It is well known that contrast normalization exists in the early visual system and boosts responsiveness at low contrast. However, in addition to firing rate, information transmission depends critically on the consistency of spike timing and spike number. It was not known how these properties would change with contrast
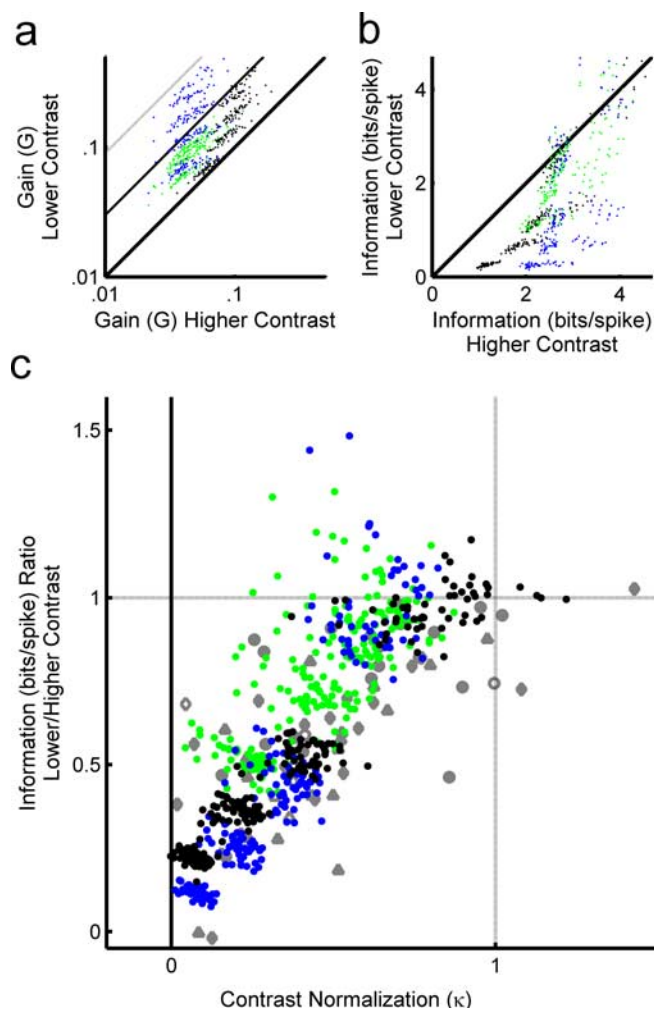
Figure 6. Contrast normalization and information preservation in a nonadapting model. Model cells with fixed, nonadapting parameters were used to simulate responses to stimuli of different contrasts. ***a***, Each symbol compares the gain of one model cell at two contrasts (compare with Fig. 3*c*). Comparisons are between 100 and 33% contrast (green symbols), 33 and 11% contrast (black symbols), or 100 and 11% contrast (blue symbols). ***b***, Mutual information in units of bits per spike (compare with Fig. 5*a*). Symbols as in ***a***. ***c***, Each symbol indicates the information ratio versus the contrast normalization, κ, for one model cell for two contrasts. Colored symbols as defined in ***a***; for comparison, results from LGN cells (from Fig. 5*c*) are shown in gray.

or whether these changes could be predicted on the basis of gain changes. We show that both spike count and spike timing are more variable as contrast decreases. These findings are consistent with the observation in retinal ganglion cells that noise is constant across contrasts (Demb et al., 2004), such that the relative amplitude of fluctuations in the membrane potential related to synaptic release (noise) becomes inconsequential for larger depolarizations. Nevertheless, many LGN cells have sub-Poisson variability and high temporal precision even at the lowest contrast tested. Average spike-timing jitter at low contrast was less than 2 ms in one-third of our cells. Therefore, reliable, precise spiking and efficient coding are not idiosyncratic to high-contrast visual stimuli.

Evidence that normalization improves information transmission

It is generally assumed that contrast normalization serves a role in making transmission of information efficient, but this need not be the case, because normalization does not take into account the

effects of contrast on timing precision and spike-count reliability. The intrinsic heterogeneity of our neural population allowed us to test this hypothesis and demonstrate that contrast normalization is correlated with the ability of the a neuron to preserve information transmission across contrasts (Fig. 5). Moreover, it is an open question whether neurons are optimized for the absolute rate of transmission (bits per second) or for the metabolic efficiency of transmission (bits per spike). We found that contrast normalization is correlated to the preservation across contrasts of information transmission in bits per spike; there was no significant correlation with the information rate in bits per second (data not shown).

Role of active adaptation questioned

It is often argued that neurons must adapt to contrast because they lack the dynamic range to discriminate both weak stimuli (low contrast) and strong stimuli (high contrast) simultaneously. Indeed, our results show that contrast normalization is associated with maintenance of neural coding efficiency, as predicted on theoretical grounds. Moreover, information transmission is better preserved than rate modulation, perhaps because of the additional changes in neural dynamics that are also associated with contrast gain control (Shapley and Victor, 1978; Victor, 1987, 1999).

Our modeling results challenge the assumption that an active process of adaptation is required for this function: at least one nonadapting model can exhibit normalization that is correlated with information preservation. Our model provides a formal summary of effects such as noise and refractoriness occurring in all of the cells and synapses from the photoreceptor to the LGN relay cell. The biophysical mechanisms underlying functional contrast normalization may occur in any or all of these cells. For example, studies investigating the cellular mechanism of contrast normalization suggest that such normalization arises in retinal bipolar cells (Rieke, 2001; Beaudoin et al., 2007), and mechanisms related to refractory periods (e.g., sodium channel inactivation) have been invoked to account for discrepancies between contrast normalization in the subthreshold and spiking responses of ganglion cells (Kim and Rieke, 2001, 2003; Zaghloul et al., 2005; Beaudoin et al., 2007).

References
Atick J, Redlich A (1992) What does the retina know about natural scenes? Neural Comput 4:196–210.

Baccus SA, Meister M (2002) Fast and slow contrast adaptation in retinal circuitry. Neuron 36:909–919.

Beaudoin DL, Borghuis BG, Demb J (2007) Cellular basis for contrast gain control over the receptive field center of mammalian retinal ganglion cells. J Neurosci 27:2636–2645.

Benardete EA, Kaplan E (1997) The receptive field of the primate P retinal ganglion cell. II. Nonlinear dynamics. Vis Neurosci 14:187–205.

Benardete EA, Kaplan E (1999) Dynamics of primate P retinal ganglion cells: responses to chromatic and achromatic stimuli. J Physiol (Lond) 519:775–790.

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

Berry MJ, Meister M (1998) Refractoriness and neural precision. J Neurosci 18:2200–2211.

Berry MJ, Warland DK, Meister M (1997) The structure and precision of retinal spike trains. Proc Natl Acad Sci USA 94:5411–5416.

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

Borst A, Flanagin VL, Sompolinsky H (2005) Adaptation without parameter change: Dynamic gain control in motion detection. Proc Natl Acad Sci USA 102:6172–6176.

Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. Neuron 26:695–702.

Brown SP, Masland RH (2001) Spatial scale and cellular substrate of contrast adaptation by retinal ganglion cells. Nat Neurosci 4:44–51.

Carandini M, Ferster D (1997) A tonic hyperpolarization underlying contrast adaptation in cat visual cortex. Science 276:949–952.

Chander D, Chichilnisky EJ (2001) Adaptation to temporal contrast in primate and salamander retina. J Neurosci 21:9904–9916.

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

Demb JB, Sterling P, Freed MA (2004) How retinal ganglion cells prevent synaptic noise from reaching the spike output. J Neurophysiol 92:2510–2519.

DeWeese M, Zador A (1998) Asymmetric dynamics in optimal variance adaptation. Neural Comput 10:1179–1202.

Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412:787–792.

Fee MS, Mitra PP, Kleinfeld D (1996) Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. J Neurosci Methods 69:175–188.

Freed MA (2005) Quantal encoding of information in a retinal ganglion cell. J Neurophysiol 94:1048–1056.

Gaudry KS, Reinagel P (2007) Contrast adaptation in a non-adapting LGN model. J Neurophysiol, in press.

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

Heitwerth J, Kern R, van Hateren JH, Egelhaaf M (2005) Motion adaptation leads to parsimonious encoding of natural optic flow by blowfly motion vision system. J Neurophysiol 94:1761–1769.

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

Jin X, Chen AH, Gong HQ, Liang PJ (2005) Information transmission rate changes of retinal ganglion cells during contrast adaptation. Brain Res 1055:156–164.

Kara P, Reinagel P, Reid RC (2000) Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. Neuron 27:635–646.

Keat J, Reinagel P, Reid RC, Meister M (2001) Predicting every spike: A model for the responses of visual neurons. Neuron 30:803–817.

Kim KJ, Rieke F (2001) Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. J Neurosci 21:287–299.

Kim KJ, Rieke F (2003) Slow Na$^+$ inactivation and variance adaptation in salamander retinal ganglion cells. J Neurosci 23:1506–1516.

Kohn A, Movshon JA (2003) Neuronal adaptation to visual motion in area MT of the macaque. Neuron 39:681–691.

Kremers J, Silveira LC, Kilavik BE (2001) Influence of contrast on the responses of marmoset lateral geniculate cells to drifting gratings. J Neurophysiol 85:235–246.

Liu RC, Tzonev S, Rebrik S, Miller KD (2001) Variability and information in a neural code of the cat lateral geniculate nucleus. J Neurophysiol 86:2789–2806.

Lu SM, Guido W, Sherman SM (1992) Effects of membrane voltage on receptive field properties of lateral geniculate neurons in the cat: contributions of the low-threshold Ca^{2+} conductance. J Neurophysiol 68:2185–2198.

Ohzawa I, Sclar G, Freeman RD (1985) Contrast gain control in the cat's visual system. J Neurophysiol 54:651–667.

Reich DS, Victor JD, Knight BW, Ozaki T, Kaplan E (1997) Response variability and timing precision of neuronal spike trains in vivo. J Neurophysiol 77:2836–2841.

Reinagel P, Reid RC (2000) Temporal coding of visual information in the thalamus. J Neurosci 20:5392–5400.

Reinagel P, Reid RC (2002) Precise firing events are conserved across neurons. J Neurosci 22:6837–6841.

Rieke F (2001) Temporal contrast adaptation in salamander bipolar cells. J Neurosci 21:9445–9454.

Ruderman DL, Bialek W (1994) Statistics of natural images: scaling in the woods. Physical Rev Lett 73:814–817.

Sanchez-Vives MV, Nowak LG, McCormick DA (2000) Cellular mechanisms of long-lasting adaptation in visual cortical neurons *in vitro*. J Neurosci 20:4286–4299.

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

Shapley R, Enroth-Cugell C (1984) Visual adaptation and retinal gain control. Prog Retinal Res 3:263–346.

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

Shapley RM, Victor JD (1981) How the contrast gain control modifies the frequency responses of cat retinal ganglion cells. J Physiol (Lond) 318:161–179.

Shou T, Li X, Zhou Y, Hu B (1996) Adaptation of visually evoked responses of relay cells in the dorsal lateral geniculate nucleus of the cat following prolonged exposure to drifting gratings. Vis Neurosci 13:605–613.

Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M (1997) Adaptation of retinal processing to image contrast and spatial scale. Nature 386:69–73.

Solomon SG, Peirce JW, Dhruv NT, Lennie P (2004) Profound contrast adaptation early in the visual pathway. Neuron 42:155–162.

Strong SP, de Ruyter van Steveninck RR, Bialek W, Koberle R (1998) On the application of information theory to neural spike trains. Pac Symp Biocomput 621–632.

Uzzell VJ, Chichilnisky EJ (2004) Precision of spike trains in primate retinal ganglion cells. J Neurophysiol 92:780–789.

van Hateren JH (1997) Processing of natural time series of intensities by the visual system of the blowfly. Vision Res 37:3407–3416.

Victor JD (1987) The dynamics of the cat retinal X cell centre. J Physiol (Lond) 386:219–246.

Victor JD (1999) Temporal aspects of neural coding in the retina and lateral geniculate. Network Comput Neural Syst 10:R1–R66.

Yu Y, Lee T (2005) Adaptive contrast gain control and information maximization. Neurocomputing 65–66:111–116.

Yu Y, Potetz B, Lee TS (2005) The role of spiking nonlinearity in contrast gain control and information transmission. Vision Res 45:583–592.

Zaghloul KA, Boahen K, Demb JB (2005) Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. J Neurosci 25:860–868.