

Temporal Coding of Visual Information in the Thalamus

Pamela Reinagel and R. Clay Reid

Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115

The amount of information a sensory neuron carries about a stimulus is directly related to response reliability. We recorded from individual neurons in the cat lateral geniculate nucleus (LGN) while presenting randomly modulated visual stimuli. The responses to repeated stimuli were reproducible, whereas the responses evoked by nonrepeated stimuli drawn from the same ensemble were variable. Stimulus-dependent information was quantified directly from the difference in entropy of these neural responses. We show that a single LGN cell can encode much more visual information than had been demonstrated previously, ranging from 15 to 102 bits/sec across our sample of cells. Information rate was correlated with the firing rate of the cell, for a consistent rate of 3.6 ± 0.6 bits/spike (mean \pm SD). This

information can primarily be attributed to the high temporal precision with which firing probability is modulated; many individual spikes were timed with better than 1 msec precision. We introduce a way to estimate the amount of information encoded in temporal patterns of firing, as distinct from the information in the time varying firing rate at any temporal resolution. Using this method, we find that temporal patterns sometimes introduce redundancy but often encode visual information. The contribution of temporal patterns ranged from -3.4 to $+25.5$ bits/sec or from -9.4 to $+24.9\%$ of the total information content of the responses.

Key words: LGN; neural coding; information theory; entropy; white noise; reliability; variability

Cells in the lateral geniculate nucleus of the thalamus (LGN) respond to spatial and temporal changes in light intensity within their receptive fields. The collective responses of many such cells constitute the input to visual cortex. All stimulus discrimination at the perceptual level must ultimately be supported by reliable differences in the neural response at the level of the LGN cell population. We are therefore interested in measuring the statistical discriminability of LGN responses elicited by different visual stimuli.

It has been shown that the LGN can respond to visual stimuli with remarkable temporal precision (Reich et al., 1997). This implies that LGN neurons have the capability to signal information at high rates. Previous estimates of the information in LGN responses have used two general approaches. The first approach, stimulus reconstruction, relies on an explicit model of what the neuron is encoding, as well as an algorithm for decoding it (Bialek et al., 1991; Rieke et al., 1997). This method has been used to place lower bounds on the information encoded by single neurons (Reinagel et al., 1999) or pairs of neurons (Dan et al., 1998) in the LGN in response to dynamic visual stimuli.

The second approach, the “direct” method, relies instead on statistical properties of the responses to different stimuli (the entropy of the responses). Because this involves only comparisons of spike trains, without reference to stimulus parameters, we need not know what features of the stimulus the cell encodes. Analysis of this type can be simplified by using a small set of stimuli and describing neural responses in terms of a few parameters, as has been done in previous studies of the LGN (Eckhorn and Pöpel, 1975; McClurkin et al., 1991).

Recently, a version of the direct method has been developed that can be applied to the detailed firing patterns of neurons in response to arbitrarily complex stimuli (Strong et al., 1998). This method provides a direct measure of how much information is contained in

a neural response, in the sense that the method is independent of any assumptions about what the neuron represents or how that information is represented. The information could be encoded at any temporal resolution and could involve any kind of temporal pattern. Here we apply this method to study the responses of individual LGN cells to a complex (high-entropy) temporal stimulus.

Because the direct method does not constrain either the temporal resolution of the code or the role of temporal patterns, the result does not by itself tell us anything about how LGN cells encode stimuli. We therefore present two further analyses. First, we explore the temporal resolution of the neural code. Second, we introduce a measure of the contribution of temporal patterns.

We distinguish three broad possibilities: (1) temporal patterns do not exist or are irrelevant to the neural code; (2) temporal patterns exist and make the neural code more redundant; or (3) temporal patterns exist and encode useful information. In our data, we find a range of results. Some cells encode information redundantly, whereas others use temporal patterns to encode visual information. In the latter case, to extract all the information from the spike trains, it would be necessary to consider temporal firing patterns; the time-varying instantaneous probability of firing would not be sufficient at any temporal resolution.

MATERIALS AND METHODS

Experimental

Surgery and preparation. Cats were initially anesthetized with ketamine HCl (20 mg/kg, i.m.) followed by sodium thiopental (20 mg/kg, i.v., supplemented as needed and continued at 2–3 mg \cdot kg⁻¹ \cdot hr⁻¹ for the duration of the experiment). The animals were then ventilated through an endotracheal tube. Electrocardiograms, electroencephalograms, temperature, and expired CO₂ were monitored continuously. Animals were paralyzed with Norcuron (0.3 mg \cdot kg⁻¹ \cdot hr⁻¹, i.v.). Eyes were refracted, fitted with appropriate contact lenses, and focused on a tangent screen. Electrodes were introduced through a 0.5 cm diameter craniotomy over the LGN. 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 Harvard Medical Area Standing Committee on Animals.

Electrical recording. Single LGN neurons in the A laminae of the LGN were recorded with plastic-coated tungsten electrodes (AM Systems, Everett, WA). In some experiments, single units were recorded with electrodes of a multielectrode array (System Eckhorn Thomas Recording, Marburg, Germany). Recorded voltage signals were amplified, filtered, and passed to a personal computer running DataWave (Longmont, CO) Discovery software, and spike times were determined to 0.1 msec resolution. Preliminary

Received Sept. 28, 1999; revised April 28, 2000; accepted May 1, 2000.

This work was supported by National Institutes of Health Grants R01-EY10115, P30-EY12196, and T32-NS07009. Christine Couture provided expert technical assistance. Sergey Yurgenson suggested the control of composing words from temporally distant bins. We thank our many colleagues who provided critical reading of this manuscript, particularly C. Brody, M. DeWeese, and M. Meister.

Correspondence should be addressed to Clay Reid, Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115. E-mail: clay_reid@hms.harvard.edu.

Copyright © 2000 Society for Neuroscience 0270-6474/00/205392-09\$15.00/0

spike discrimination was done during the experiment, but analysis is based on a more rigorous spike sorting from postprocessing of the recorded waveforms. The power spectra of the spike trains did not contain a peak at the stimulus frame rate (128 Hz).

For the purposes of this analysis, it is crucial that any trial-to-trial variability in the recorded response be unequivocally attributable to neural noise rather than noise introduced at the level of data acquisition. Therefore, we report here only the results from extremely well isolated single units, with strict absolute refractory periods and spike amplitudes several SDs above the noise. Voltage traces surrounding each spike were examined for spike shape, as well as for precision of recorded spike times. Of the 27 cells recorded, 13 cells were excluded for imperfect spike isolation, and one cell was excluded because it responded poorly (<1 spike/sec) to our stimulus. The remaining recordings we believe to be entirely free of errors in the form of missed spikes, extraneous spikes, or spike time jitter greater than 0.1 msec. Furthermore, we used Monte Carlo simulations to simulate the effects of contamination with such errors, had they been present. We showed that if 2% of spikes were missed, 2% extraneous events were recorded as spikes, or 2% of spike times were jittered by the width of a spike, our information estimate would have decreased by 1.4, 2.8, or 0.1%, respectively.

We define a “burst” as a group of action potentials, each of which is <4 msec apart, preceded by a period of >100 msec without spiking activity. This criterion was shown previously to reliably identify bursts that are caused by low-threshold calcium spikes in LGN relay cells of the cat (Lu et al., 1992). We define the “burst fraction” of a spike train as the fraction of all spikes that were part of a burst.

Visual stimuli. A spatially uniform (full-field) white-noise visual stimulus was constructed by randomly selecting a new luminance every 7.8 msec (Fig. 1*a*, *stim*). The luminance values were drawn from a natural probability distribution (van Hateren, 1997), but the stimulus was otherwise artificial and had no temporal correlation between frames. The resulting stimulus had an entropy of 7.2 bits/frame [$\sum_i P(i) \log_2 P(i)$, where $P(i)$ is the probability of luminance i], for a rate of 924.4 bits/sec, which well exceeds the coding capacity (i.e., the measured entropy) of individual neural responses. The intensity distribution was dominated by low intensities, such that luminance was $<33\%$ of the maximum in 73% of frames, and changed by $<33\%$ of the total range in 77% of transitions. This produced moderate contrast compared with, for instance, a binary black-to-white flicker (variance, 0.04 vs 0.25).

We constructed a single 20-min-long stimulus consisting of 128 repeats of a single 8 sec sample interleaved with 32 unique 8 sec samples. We presented this 20 min composite stimulus in a single continuous trial. Most cells responded well to this modulated full-field stimulus. Stimuli were presented on a computer monitor controlled by a personal computer with an AT-Vista graphics card, at 128 Hz and 8 bit gray scale, at a photopic mean luminance.

Numeric calculations

We measured the visual information in neural responses by measuring entropy rates from probability distributions of neural responses, as explained in Results (Fig. 2). First, neural responses were represented as binned spike trains in which the value of a time bin was equal to the number of spikes that occurred during that time interval. Then we analyzed short strings of bins, or words. This representation depends on two parameters, which we varied: the size of the time bins, $\delta\tau$, and the number of bins in the words, L . We measured two forms of response entropy, the noise entropy, H_{noise} , which reflects the trial-to-trial variability of responses to a repeated stimulus, and the total entropy, H_{total} , which reflects the variability of responses to all (nonrepeated) stimuli in the ensemble. H_{noise} was calculated from the distribution of responses at a fixed time t relative to stimulus onset, in 128 repeats of the same sample of the white-noise stimulus. We performed a separate calculation of H_{noise} for many different values of t (separated by one bin) within the 8 sec stimulus. We then averaged the result over all values of t to get the average noise entropy. H_{total} was calculated in the identical manner. We estimated the entropy from the distribution of responses over responses to 256 unique stimuli at each time t and then averaged the result over t .

Data adequacy for entropy calculations. In any analysis of this kind, misleading results could be obtained if the amount of data were insufficient for estimating the probabilities of each response. In pilot studies with 1024 repeats, we estimated that 128 repeats were adequate for measuring the entropy of repeated stimuli. We used twice as many samples for H_{total} as for H_{noise} to compensate for the approximately twofold difference in entropy. For every entropy we computed, 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). The correction is obtained by fitting a second-order polynomial to $1/(\text{fraction of data})$ versus estimate. We imposed two criteria to ensure data adequacy. First, we required that the total correction for finite data size was $<10\%$. Second, we required that the second-order term of this correction was negligible, $<1\%$. Two cells were discarded from analysis because of data inadequacy by these criteria.

We extrapolated to infinite word length from the linear part of the curve H versus $1/L$ (see Results) (Fig. 2*c*) as described by Strong et al. (1998). Specifically, we found the point of minimum fractional change in slope in

each curve and used the four values of L up to and including this point for extrapolation to infinite L .

Similar information rates and pattern effects were obtained using fewer or more samples to estimate H_{total} and using several alternative criteria for data adequacy (data not shown).

RESULTS

Neural responses to a dynamic visual stimulus

We recorded from individual neurons in the LGN of anesthetized cats while presenting spatially uniform (full-field) visual stimuli with a random time-varying luminance. We repeated one 8 sec sample of this white-noise stimulus 128 times, interleaved among 256 sec of unique samples. We report results from 11 exceptionally well isolated cells. Results from one example cell are shown in Figures 1–4.

The responses to different stimulus samples were highly variable (Fig. 1*a*, *unique*), whereas the responses to any repeated sample were highly reliable (Fig. 1*a*, *repeat*). From the peristimulus time histogram (PSTH) (Fig. 1*a*, *PSTH*), it is apparent that any given stimulus evoked very precise responses. When examined at a fine scale, many of the peaks in the PSTH have widths of ~ 1 msec (Fig. 1*b*) ($\sigma = 0.6$ msec). This is somewhat higher than the temporal precision seen by Reich and colleagues in LGN responses to high-contrast drifting gratings ($\sigma = 5$ msec) (Reich et al., 1997) or the precision reported in the retina of other species (Berry et al., 1997). To explore the implications of this precise timing for visual coding, we turned to an information theoretic approach.

Information measured directly from spike train entropy

A neural response encodes information when there is a narrow distribution of responses to any given stimulus when it is repeated (Fig. 1*a*, *repeat*) compared with the distribution of possible responses to all stimuli (Fig. 1*a*, *unique*). The variability underlying these probability distributions can be quantified by their entropy (Shannon, 1948). The total entropy, H_{total} , is measured from the responses to unique stimuli and reflects the range of responses used for representing the entire stimulus ensemble. The total entropy sets an upper bound on the amount of visual information that could be encoded by the responses. This limit is reached only if there is no noise, meaning that the response to any given stimulus is deterministic. The trial-to-trial variability in the responses to a repeated stimulus is given by the noise entropy, H_{noise} . In general, the visual information in the response (I) can be measured by subtracting the average noise entropy from the total entropy (DeWeese, 1996; Zador, 1996; Strong et al., 1998):

$$I = H_{\text{total}} - H_{\text{noise}} \quad (1)$$

Throughout this paper, we divide both entropy (H) and information (I) estimates by duration in time, such that all estimates are reported as entropy or information rates in units of bits per sec.

To compute these quantities, we represented neural responses as binned spike trains in which the value of a time bin was equal to the number of spikes that occurred during that time interval. We then considered the probability distributions of short strings of bins, or words. The resulting estimate of entropy depends on two parameters: the temporal resolution of the bins in seconds, $\delta\tau$, and the number of consecutive bins in our words, L (Fig. 2*d*) (Strong et al., 1998):

$$H(L, \delta\tau) = - \frac{1}{L\delta\tau} \sum_{w \in W(L, \delta\tau)} P(w) \log_2 P(w) \quad (2)$$

where w is a specific word (spike pattern), $W(L, \delta\tau)$ is the set of all possible words comprised of L bins of width $\delta\tau$, and $P(w)$ is the probability of observing pattern w in a set of observations.

For any given word length L and resolution $\delta\tau$, we counted the number of times each distinct word occurred in the responses, for either the total stimulus ensemble (Fig. 2*a*, H_{total}) or repeated stimulus samples (Fig. 2*b*, H_{noise}). From the corresponding probability distributions, we estimated the entropy of the response (Eq.

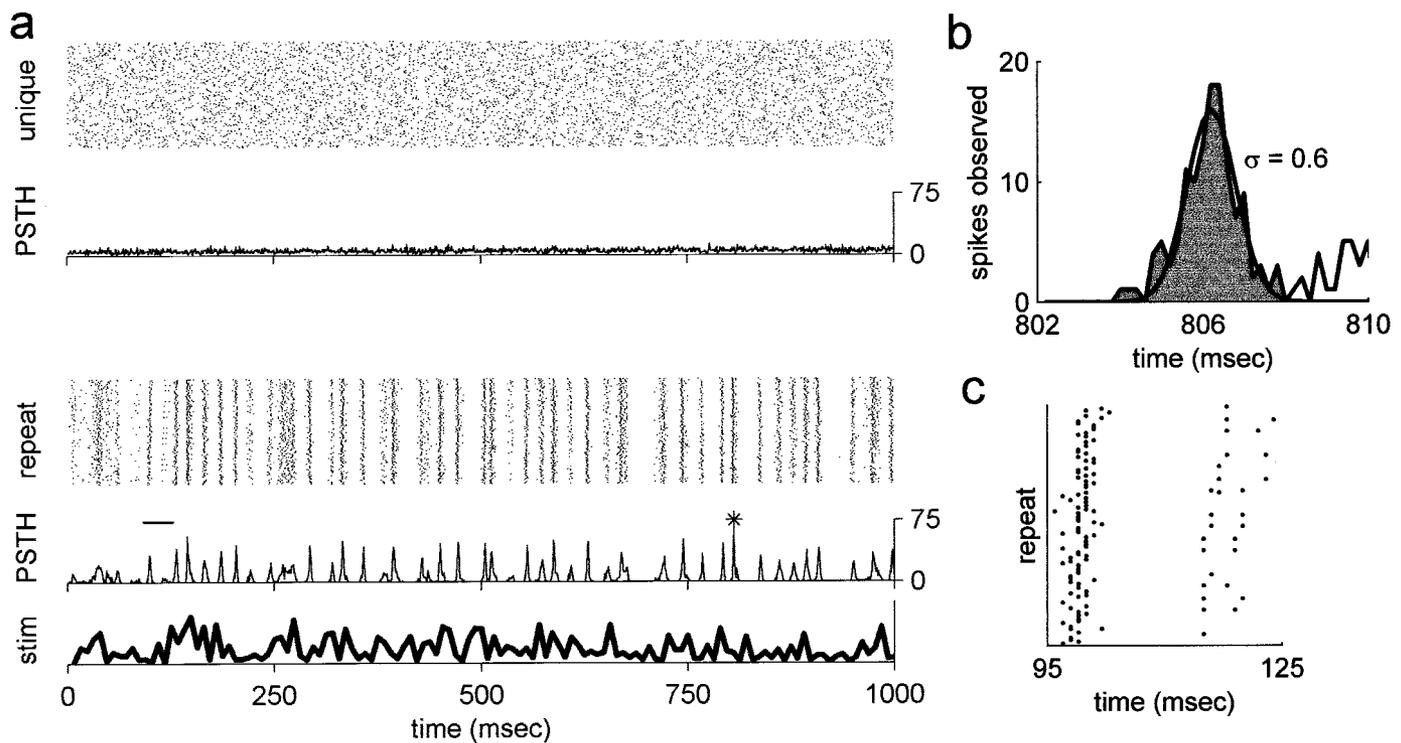


Figure 1. LGN responses to nonrepeated and repeated white-noise stimuli. Responses from an ON center X cell are shown (the cell in our sample that responded best to the full-field stimulus). *a*, Responses to 128 unique samples of the visual stimulus are shown at the top as rasters (*unique*) in which each row represents a different 1 sec segment of the neural response, and each dot represents the time of occurrence of a spike. Below the rasters is a PSTH from the 128 rasters, in which the value for each 1 msec bin is defined as the number of times a spike was observed in that bin in the 128 trials shown. The second set of rasters (*repeat*) and corresponding PSTH are from 128 repeats of one particular 1 sec sample of the white-noise visual stimulus. In the experiment, *repeat* and *unique* samples were interleaved but occurred in the order shown. The luminance time course of the stimulus that corresponds to the *repeat* rasters and PSTH is plotted at the bottom. The peak expanded in *b* is marked with an asterisk. The rasters expanded in *c* are marked by the bar. *b*, A narrow peak from PSTH in *a*, expanded and binned at 0.2 msec resolution. The peak defined by the shaded area (126 spikes in 128 repeats) was best fit by a Gaussian of $\sigma = 0.60$ msec, also shown. *c*, Expanded rasters chosen to illustrate noise. In the first half of this window, 95 of 128 trials contained a single spike, but only two trials contained two spikes. In the second half of the window, only 17 trials contained any spikes but then usually a pair of spikes ~ 5 msec apart. Finally, 16 of the 17 trials containing any spikes in the second half of the window lacked spikes in the first half.

2). Assuming sufficient data, this method sets an upper bound on the entropy rate. The true entropy rate is reached in the limit of sufficiently small time bins and infinitely long words, which we estimated by extrapolating to infinite L (Eq. 2, $\lim L \rightarrow \infty$, $\delta\tau = 0.6$ msec) (Strong et al., 1998), as illustrated in Figure 2c. From the extrapolated entropy rates and Equation 1, we obtained our best estimate of the mutual information between the spike train and the stimulus.

According to this method, the responses shown in Figure 1 contain 102 bits/sec of visual information about this stimulus set, which is about an order of magnitude higher than had been demonstrated previously in the LGN (Eckhorn and Pöpel, 1975; McClurkin et al., 1991; Dan et al., 1998; Reinagel et al., 1999). The meaning of 102 bits/sec is that the spike train contains as much information as would be required to perfectly discriminate 2^{102} different 1-sec-long samples of this white-noise stimulus. Another way to state the same finding is that this neural response contains enough information to support a reliable binary stimulus discrimination approximately every 10 msec. This conclusion is free of assumptions about the nature of the neural code and consequently tells us little about that code. To explore how this information is represented, we went on to compare the information rate to the estimates obtained with varying bin size and word length.

Temporal resolution of visual information

First, we asked how our estimate of encoded information depended on the temporal resolution of analysis, $\delta\tau$. At each temporal resolution, we computed several estimates as a function of word length L (Eqs. 1, 2) and then extrapolated to infinitely long words. In long time bins ($\delta\tau > 64$), the number of spikes was almost constant and

therefore carried little information about the dynamic stimulus. Our estimate of the visual information increased with the temporal resolution of the analysis (Fig. 3). The estimate continued to increase with temporal precision to the smallest bin size we analyzed, 0.6 msec, which is consistent with the high temporal precision evident in the PSTH (Fig. 1). To account for all of the information in the response (102 bits/sec), a resolution of at least $\delta\tau = 0.6$ msec was required, which is well below the refractory period of the neuron (2.7 msec) (Fig. 3, *a*). This implies that the timing of a single spike can be much more precise than the smallest interval between spikes.

Measuring the effect of temporal structure on coding

The previous analysis allowed for the possibility of any temporal structure within the spike train. There are many known physiological properties that would result in temporal patterns in the response, such as absolute and relative refractory periods and the presence of stereotyped bursts. Such patterns result in statistical dependence between time bins. However, it is not obvious what role these patterns play in coding visual information.

The probability of firing clearly varies with time (Fig. 1*a*, *PSTH*). We would neglect any additional temporal structure in the spike train if we simply considered the information in the average single time bin (Eqs. 1, 2, $L = 1$), instead of using long words. The information rate estimated from $L = 1$ represents an approximation on the assumption of independence between bins. If each bin were indeed independent of other bins (that is, if there were no further temporal interactions between bins beyond those reflected in the *PSTH*), then the information rate calculated at any word length L would be equal to that calculated with $L = 1$.

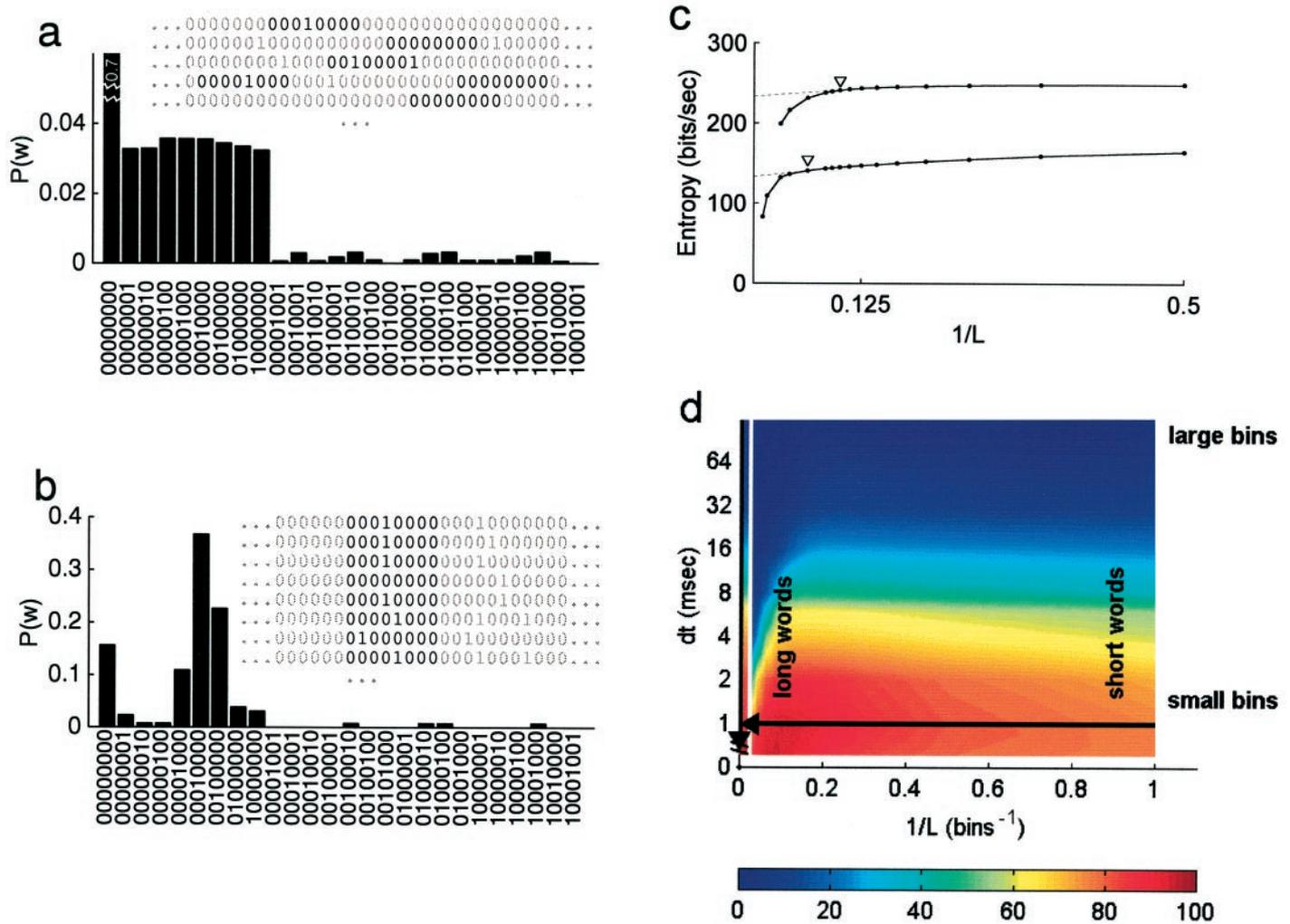


Figure 2. Illustration of calculation of mutual information. Data are same as shown in Figure 1. *a*, Probability distribution of words of length $L = 8$, binned at $\delta\tau = 1$, for the calculation of H_{total} . $P(w)$ for $w = 00000000$ is far off scale (0.70). Although 2^8 patterns were possible, only 26 actually occurred ($n = 31776$). *Inset*, Samples from five responses to nonrepeated stimuli (Fig. 1*a*, *unique*). Several eight bin words are highlighted. *b*, Probability distribution of words of length $L = 8$, binned at $\delta\tau = 1$, for the calculation of H_{noise} at one particular word position. Over 128 repeats, only 13 patterns were observed at this position of the 26 patterns observed for the entire stimulus ensemble (*a*). *Inset*, Samples from eight of the 128 responses to the repeated stimulus (Fig. 1*a*, *repeat*). A particular eight bin word is highlighted, which corresponds to a fixed time in the repeated stimulus. *c*, Estimated entropy rate of the responses, H (Eq. 2), is plotted against the reciprocal of word length, $1/L$. Note that longer words are to the left in this plot. Both H_{total} (*top*) and H_{noise} (*bottom*) decrease gradually with increasing L , as expected if there are any correlations between bins. For very long words, H_{noise} and H_{total} fall off catastrophically, which indicates that there are not enough data for the calculation beyond this point. *Dashed lines* show the extrapolations from the linear part of these curves to infinitely long words ($\lim L \rightarrow \infty$), as described by Strong et al. (1998). The point of least fractional change in slope was used as the maximum word length L (*arrows*) used for extrapolations. Mutual information I is the difference between these two curves (Eq. 1). *d*, Parameter space of the calculation. I is estimated over a range of L (plotted as $1/L$, horizontal axis) and a range of $\delta\tau$ (vertical axis). The resulting estimate $I(L, \delta\tau)$ obtained with different parameter values is indicated by *color* (interpolated from discrete samples). Values to the left of the gap reflect extrapolations to infinite word length ($\lim L \rightarrow \infty$, i.e., $1/L \rightarrow 0$). *Arrows* indicate slices through parameter space: $L = \infty$ (*vertical arrow*, replotted in Fig. 3) and $\delta\tau = 1$ (*horizontal arrow*, replotted in Fig. 4). Point at origin indicates the true information rate, which is obtained in the limit $L \rightarrow \infty$ at sufficiently small $\delta\tau$. (With finite data, the estimate is not well behaved in the limit of $\delta\tau = 0$).

We introduce a new quantity Z for the correction in our information estimate when we take temporal patterns into account compared with the estimate we would have obtained if we had ignored them:

$$Z(\delta\tau) \equiv \lim_{L \rightarrow \infty} I(L, \delta\tau) - I(L = 1, \delta\tau) \quad (3)$$

For a discussion of how Z relates to other quantities, see Appendix.

If temporal patterns as such encode visual information, then the pattern correction term Z is positive, because one time bin is synergistic with other bins in encoding information. Conversely, if temporal patterns merely introduce redundancy into the neural code, the pattern correction Z is negative, because the stimulus information conveyed by one time bin is partly redundant with the information contained in other bins.

For the responses shown in Figures 1–3, the pattern correction Z

was positive. The estimate obtained at 1 msec resolution but neglecting patterns, $I(L = 1, \delta\tau = 1)$, was 79 bits/sec, whereas our best estimate of the true information at this resolution, $I(\lim L \rightarrow \infty, \delta\tau = 1)$ was 97 bits/sec (Fig. 3) for a pattern correction $Z(\delta\tau = 1) = +18$ bits/sec. Thus, not only is information encoded in spike timing at 1 msec resolution, but extended patterns of spikes must be considered to obtain all of this information. At even higher temporal resolution, $Z(\delta\tau = 0.6) = 25$ bits/sec, which indicates that patterns accounted for 25 of the 102 bits/sec total information. The strength of this analysis is that the information we attribute to spike patterns (Z) is completely distinct from the information attributable to the temporal precision of firing-rate modulation (PSTH).

As we increased the word length L at $\delta\tau = 1$ msec, our estimate of the mutual information increased gradually (Fig. 4, *filled symbols*). Using a lower resolution analysis to explore longer time windows ($L < 10$, $\delta\tau > 1$ msec), we found evidence for pattern

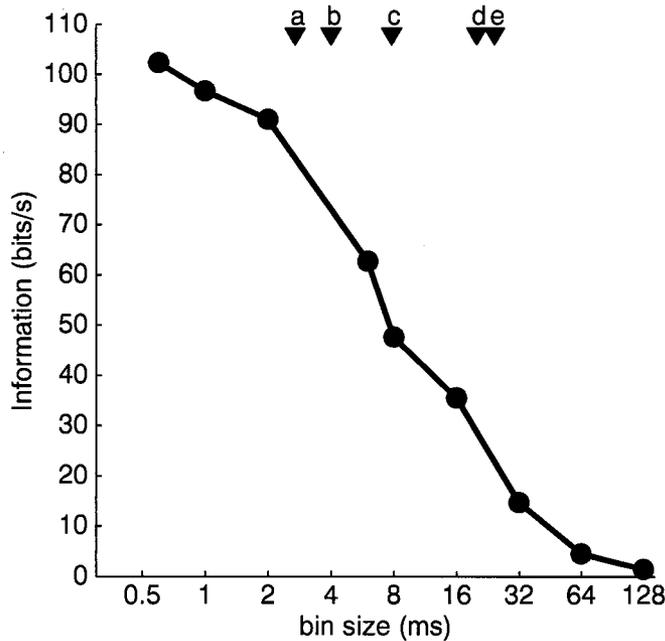


Figure 3. Temporal resolution of visual information. For each bin size $\delta\tau$ (horizontal axis), we computed the mutual information between stimulus and response (Eq. 1) from the distribution of observed firing patterns (Eq. 2), over a range of word lengths L . Data are from Figure 2*d*. Results shown are the extrapolations to $\lim L \rightarrow \infty$ at each temporal resolution. Arrowheads indicate notable time scales in the data (see Discussion): *a*, absolute refractory period (2.7 msec); *b*, mode interspike interval (4.0 msec); *c*, stimulus frame rate (7.8 msec); *d*, duration of main peak of impulse response of the neuron (20.0 msec); *e*, mean interspike interval (24.4 msec).

information up to at least 64 msec (data not shown). We note that even short words often contain more than one spike. At this temporal resolution ($\delta\tau = 1$ msec), there were no words with multiple spikes up to $L = 3$ because of the refractory period of the cell. For longer words ($L \geq 4$), the fraction of nonempty words that contained multiple spikes increased linearly with word length, crossing 12.5% at $L = 10$.

We suggest that the information in firing patterns is the result of temporal structure within the spike trains that occur on a short time scale. To test this, we performed a control in which words were composed not from L contiguous bins but from L bins well separated in time (chosen randomly from 8 sec samples). To compute H_{noise} , a set of words was extracted using the same bins from each repeat. In these words of nonadjacent bins, our estimate of the information rate depended only very weakly on word length (Fig. 4, *scrambled*). This confirmed that the relevant temporal interactions are local in time.

Population results

Among the 11 cells we analyzed, we obtained a range of firing rates in response to this full-field white-noise stimulus. Absolute information rate was strongly correlated with firing rate, such that the information encoded per spike was relatively constant at 3–4 bits/spike (Fig. 5*a*). Coding efficiency can be defined as the fraction of response entropy at a given bin size $\delta\tau$ that is used to carry stimulus information (Rieke et al., 1997). A perfectly noise-free code would have an efficiency of 1. Based on our analysis, we estimated the coding efficiency as I/H_{total} , computed at $\lim L \rightarrow \infty$, $\delta\tau = 1$. Over our population of cells, estimated coding efficiency was 0.53 ± 0.10 (mean \pm SD).

All cells showed a precision of spike timing of 2 msec or better. The estimate obtained at $\delta\tau = 1$ msec was 6% higher than the estimate at $\delta\tau = 2$ msec in the example shown in Figure 3. The estimate at $\delta\tau = 1$ msec was higher than at $\delta\tau = 2$ msec for 9 of 11 cells (increasing by $10 \pm 2\%$). For the remaining 2 of 11 cells, estimates at $\delta\tau = 1$ and 2 msec differed by $<1\%$, but estimates at

$\delta\tau = 2$ msec were higher than estimates at $\delta\tau = 4$ msec (by $23 \pm 9\%$).

Not all cells encoded information in spike patterns. Analyzed at 1 msec resolution, the pattern correction Z ranged from -9.4 to $+18.6\%$, or -3.4 to $+18.0$ bits/sec. The pattern correction was correlated with the firing rate of the response, but this trend was not statistically significant ($p = 0.07$) (Fig. 5*b*). For Figures 1–4, we selected a cell with a particularly high firing rate and a large, positive Z to demonstrate a strong example of information in spike patterns.

One of our cells showed a substantial net redundancy in spike patterns, with a pattern correction of -9.4% or -3.4 bits/sec. Interestingly, this cell had a high frequency of bursting, a firing pattern characteristic of LGN cells (Sherman, 1996). In our sample, the frequency of bursting ranged from 0 to 15% (see Materials and Methods) and was inversely correlated with Z ($p < 0.005$) (Fig. 5*c*).

Simulations

To gain insight into what kinds of neural mechanisms could produce either 0, negative, or positive Z , we analyzed three kinds of artificial spike trains. In condition A, spike trains were generated by a Poisson process according to a time-varying firing rate taken from that of a real neuron (Fig. 1*a*, *PSTH*). Thus firing probability was modulated with the same high temporal precision as the cell. However, the estimate of information did not vary with word length L , such that $Z = 0$ (Fig. 4, *Poisson*).

In condition B, spike trains were generated exactly as in A, but for every spike generated in A, a doublet of spikes several milliseconds apart was produced in B. In this case, separate 1 msec time bins carried partly redundant information, which would be “counted twice” if bins were assumed to be independent. Thus, we predicted that this kind of temporal structure would add redundancy to the neural code. Indeed, the estimate of information using $L = 1$ was an almost twofold overestimate of the true information (compared with $\lim L \rightarrow \infty$). Therefore, the pattern correction was negative: $Z = -44\%$. This model suggests a simple explanation for

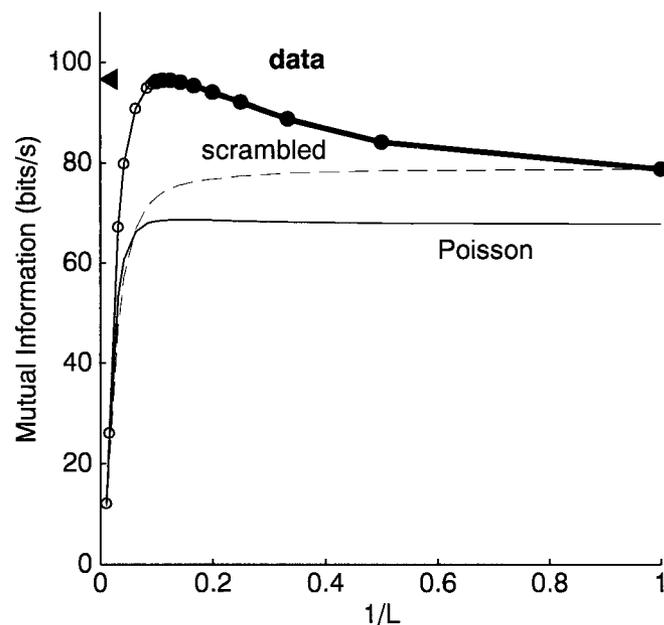


Figure 4. Effect of word length on estimated visual information. Information I between the neural response and the visual stimulus (Eq. 1) as a function of the reciprocal of word length $1/L$, computed at temporal resolution $\delta\tau = 1.0$ msec. Data are from Figure 2*d*. Filled circles, Word lengths that satisfied our data adequacy criterion (see Materials and Methods). Open circles, Word lengths that failed our criterion (insufficient data). Dashed line, *scrambled*, Words composed of bins far apart in time (chosen randomly from throughout 8 sec samples). Thin solid line, *Poisson*, Results for a Poisson model spike train whose time-varying rate matched the PSTH of this cell (Fig. 1*a*).

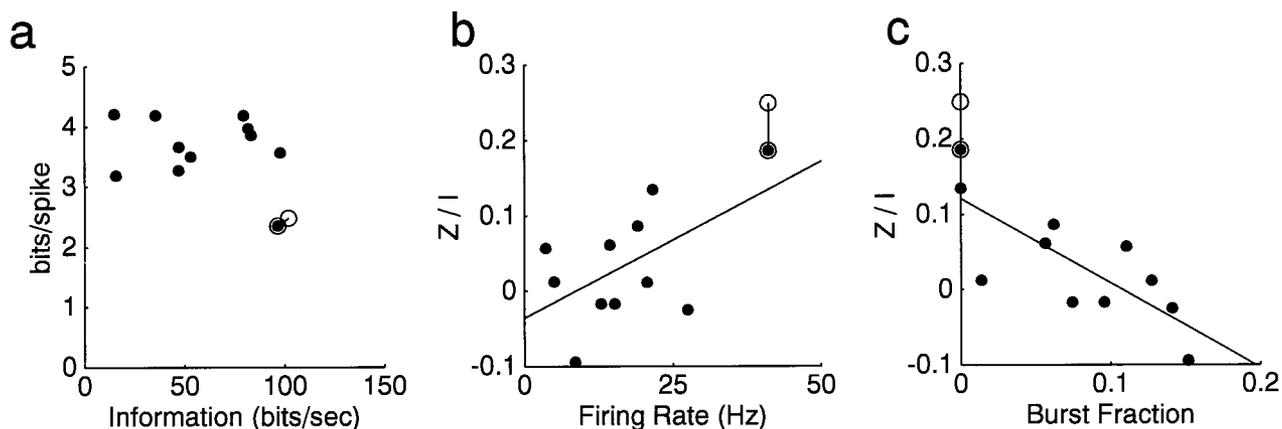


Figure 5. Summary of results for all cells in this study. Each filled circle represents one cell. Results are shown for $\lim L \rightarrow \infty$ and $\delta\tau = 1.0$ msec. Circled symbol is the cell shown in Figures 1–4. Open circle is that cell analyzed at $\delta\tau = 0.6$ msec. *a*, Mutual information, I , in units of bits per spike (vertical axis) versus units of bits per seconds (horizontal axis). *b*, Pattern correction Z , expressed as a fraction of total information, $Z/I(\lim L \rightarrow \infty)$, is plotted as a function of the average firing rate of the response of the cell. Line shown is least square fit to data: slope, $+0.42\%$ per Hertz. This trend is not statistically significant ($R^2 = 0.31$, $p = 0.07$). *c*, Pattern correction Z plotted as a function of the burst fraction of the spike train (see Materials and Methods). Line shown is least square fit to data: slope, -1.1% per percent bursting. This relationship is highly statistically significant ($R^2 = 0.60$, $p < 0.005$).

why cells with stereotyped bursts showed temporal redundancy (low or negative Z).

Finally, in condition C, spike trains were generated as in A, but noise was added in the form of spike time jitter. Nearby spikes were jittered by a correlated amount, simulating a low-frequency noise source. The result was that the precision of individual spikes was degraded, but local temporal patterns were preserved. In this case, the noise (temporal jitter) would be counted twice if bins were assumed to be independent ($L = 1$). Thus, we predicted that responses would be more reproducible at the level of words than one would expect from the variability of each bin. Simulations confirmed that, in spike trains of this kind, $L = 1$ produced an underestimate of the true visual information. Therefore, the pattern correction was positive. For example, one such model with a $1/f$ noise spectrum and a spike jitter of $\sigma = 2.5$ msec had a pattern correction of $Z = +14.4\%$.

In our data, it is not uncommon to observe jitter from trial to trial in which the relationships between nearby spikes are preserved. Model C suggests that this would be one source of temporal coding in our data, but there are probably others. Another source of trial-to-trial noise is the variable presence of a spike as opposed to jitter in the timing of the spike. We sometimes observe positive correlations between two unreliable events; if one spike is present, both are. Conversely, we find cases of “either–or” firing in which events separated by several milliseconds in time occur in a mutually exclusive manner. We suspect that these types of temporally correlated noise also contribute to pattern coding. Figure 1c illustrates all three types of noise in our data.

DISCUSSION

We have demonstrated information rates in excess of 100 bits/sec and rates of 3–4 bits/spike in spike trains of single neurons in the cat LGN. Thus, LGN cells can carry high rates of information about dynamic stimuli, as a similar analysis has revealed of several other types of neurons (Berry et al., 1997; de Ruyter van Steveninck et al., 1997; Buračas et al., 1998; Strong et al., 1998). Previous studies of the LGN have demonstrated much lower information rates (Eckhorn and Pöpel, 1975; McClurkin et al., 1991; Dan et al., 1998; Reinagel et al., 1999). We attribute the difference to two factors. First, the previous studies made use of approximations that were known to underestimate the information content, whereas our direct method was much less constrained. Second, our full-field white-noise stimulus may have been a stronger stimulus, at least for some cells, than the stimuli used previously.

One of the fundamental questions of neuroscience is how sensory information is encoded by neural activity. The first evidence

on this question was the discovery that the amount of pressure of a tactile stimulus correlated with the number of spikes elicited in a somatosensory neuron over the duration of the stimulus (Adrian and Zotterman, 1926). Thus, firing rate could be said to encode that stimulus parameter. Most of our present understanding of sensory neural coding is based on similar observations; the firing rates of different neurons are correlated with different stimulus parameters (Barlow, 1972).

It has long been recognized that such a “rate code” may not exhaustively describe the neural code and that additional features of spike timing may also carry stimulus information (MacKay and McCulloch, 1952). In the recent literature on neural coding, there has been considerable debate on the existence and nature of such “temporal codes” (Shadlen and Newsome, 1994, 1995, 1998; Ferster and Spruston, 1995; Sejnowski, 1995; Stevens and Zador, 1995; Rieke et al., 1997; Victor, 1999). This debate has been fueled by accumulating evidence that precise times of spikes do carry information, especially when stimuli themselves are temporally varying.

Much of the discussion in the literature has confounded two different ideas. The first idea is that the firing rate (or probability of firing) may vary on a much finer time scale than was classically explored: milliseconds rather than hundreds of milliseconds. The second idea is that temporal patterns of firing must be considered to fully describe the neural code. To date, most of the evidence for information in spike timing in single neurons has been consistent with the first idea. We have proposed a way to uncouple these two phenomena and have demonstrated that both forms of coding are evident in our data.

Our results constitute an existence proof that, in the LGN, visual information can be encoded both by temporally precise modulation of firing rate, to ~ 1 msec resolution, and also in firing patterns that extend over time. The time scales and relative contributions of spike timing and spike pattern will presumably be different for different types of neurons. From our small sample of cells, we cannot yet say whether there will be systematic differences between different cell types in the LGN. Even for a single neuron, results may depend significantly on the particular set of stimuli used. Finally, we have not considered spike patterns across multiple cells, which may introduce redundancy or may encode additional visual information (Abeles et al., 1994; Meister et al., 1995; Funke and Wörgötter, 1997; Dan et al., 1998).

Temporal precision of firing

We began by analyzing the temporal precision with which spike timing is visually driven. Sharp peaks in the PSTH revealed that rate was modulated on time scales as fine as 1 msec (Fig. 1). As a

result of this precision, our estimates of visual information depended strongly on the temporal resolution of the analysis, down to at least 1 msec (Fig. 3).

Similar studies have found high temporal precision in several other visual cell types (Berry et al., 1997; de Ruyter van Steveninck et al., 1997; Buračas et al., 1998; Strong et al., 1998) (but see Dimitrov and Miller, 2000). The precision we observe is unusually high, but direct comparisons are difficult because the temporal precision depends on the stimulus set used. For example, when cat LGN cells were studied with drifting sinusoidal gratings, a somewhat lower precision was found in responses to high-contrast gratings ($\sigma = 5$ msec), and substantially less precision was seen in responses to moderate contrast gratings (Reich et al., 1997). We also observed lower precision in the responses to checkerboard M-sequence stimuli (Reid et al., 1997) (~ 10 msec, our unpublished observation). The white-noise stimulus used here had moderate contrast (Fig. 1*a*) (see Materials and Methods). It remains possible that other stimuli could elicit even higher temporal precision in LGN responses.

Previous discussions have proposed criteria for distinguishing a rate code from a temporal code on the basis of the time scale alone. For example, when long time scales are considered, firing rate can be defined by the number of events counted from an individual cell in a single trial. When very short time scales are used, firing probability would have to be estimated from many trials or cells. Thus, one relevant time scale is the mean interspike interval; time bins longer than this have on average >1 spike and shorter bins have on average <1 spike (Fig. 3, *e*). However, we note that the mode interval (Fig. 3, *b*) is typically far shorter than the mean interval. The absolute refractory period of the cell (Fig. 3, *a*) defines the absolute minimum interval and places a limit on the maximum firing rate the neuron could sustain.

Another intrinsic time scale in the experiment is the integration time of the neuron. For instance, this might be represented by the duration of the positive phase of the neural response to a brief stimulus pulse (Fig. 3, *d*). A convolution of the stimulus by the impulse response would be expected to temporally filter (blur) the stimulus at approximately this time scale.

Alternatively, one might compare the precision of the neural response with the time scale of the stimulus rather than any neural parameter. It has been argued that, when temporal structure in the response matches the temporal structure in the stimulus, the neuron simply represents time with time. According to this view, if the spike train encodes information using a higher temporal precision than the time scale on which the stimulus changes, the neuron would be using temporal patterns to encode something nontemporal (Theunissen and Miller, 1995). For time-varying stimuli, a dividing line might then be drawn at the maximum temporal frequency of the stimulus. In a stimulus like ours, this would typically be given as the frame update rate (Fig. 3, *c*), although we note that the physical stimulus contains much higher temporal frequencies. For transiently presented, static stimuli, this time scale corresponds to the stimulus duration (McClurkin et al., 1991; Heller et al., 1995; Victor and Purpura, 1996). A limiting case for this distinction, not addressed in our study, is to compare time-varying stimuli with constant stimuli (de Ruyter van Steveninck et al., 1997; Buračas et al., 1998; Warzecha et al., 1998).

We believe that all of these arguments identify landmarks along a continuum and that information encoded in the firing rate of a cell at any time scale is a direct extension of the idea of a rate code. A qualitative distinction can be made between all of these forms of coding and any code that depends on relationships within the spike train. Such pattern information cannot be extracted from the rate (or probability) of firing alone, regardless of time scale.

Poisson models

A common model for a neuronal rate code is a homogeneous Poisson process whose mean rate is stimulus-dependent (Shadlen and Newsome, 1994, 1995). A natural extension of this model is an inhomogeneous Poisson process whose time-varying mean rate is

determined by the stimulus (Buračas et al., 1998). We constructed model Poisson spike trains using the observed PSTH for the time varying rate and found that this model encoded 30% less visual information than the real data (de Ruyter van Steveninck et al., 1997). However, even when we neglect temporal patterns (using $L = 1$, $\delta\tau = 1$ msec), the model Poisson trains had higher estimated entropy and less estimated information than the real spike trains (Fig. 4, *Poisson*). Therefore, even if the cell had had no pattern information ($Z = 0$), it would have encoded more information than this Poisson control. Extrapolated to infinite word length, the difference between the cell and the Poisson model was 29 bits/sec, but only approximately half of this discrepancy ($Z = 18$ bits/sec) was directly attributed to temporal patterns by our analysis.

The information estimates from the cell and the Poisson model differ at $L = 1$ because the number of spikes in each time bin was more precise than expected from a Poisson process, as has been observed for several other types of neurons (Levine et al., 1988; Berry et al., 1997; de Ruyter van Steveninck et al., 1997). The variance of spike count divided by the mean spike count was 0.91 ± 0.04 in 1 msec bins and 0.38 ± 0.23 in 128 msec bins ($n = 11$ cells) compared with a ratio of 1 for any Poisson process. As a result, Poisson models underestimated the information rates of each of our cells (by $19.9 \pm 9.6\%$, $n = 11$, at $\delta\tau = 1$ msec) whether the cells had negative or positive Z . We conclude that a comparison between the Poisson control and the actual data are not sufficient to show that information is coded in temporal patterns of firing.

Temporal patterns of firing

In our second analysis, we held the temporal resolution of the analysis constant and considered the role of temporal patterns in the neural code. We define a quantity, Z , that is positive when information is encoded in temporal patterns and negative when patterns constitute redundancy in the code. We found that as much as 25% of the visual information could be attributed explicitly to firing patterns. Using 1 msec bins, we had enough data to explore words of up to 10 msec, and information continued to increase with L up to this limit (Fig. 4). Using lower resolutions to explore longer time windows, we estimated that patterns contain information to at least 64 msec. Interestingly, a quite different analysis revealed a similar time scale for temporal patterns in cortical responses to low-entropy stimuli (periodic stimuli or transiently presented, static textures) (Victor and Purpura, 1997; Reich et al., 1998).

Temporal patterns could be caused by a direct effect of one stimulus-evoked spike on the subsequent ability of the stimulus to evoke other spikes at nearby times in the same trial. Examples of this type that have been discussed in the context of information coding include refractoriness (Berry and Meister, 1998) and some forms of bursting (Cattaneo et al., 1981; DeBusk et al., 1997; but see below).

Alternatively or in addition, the ability of the stimulus to evoke spikes could be modulated by an external noise source that has a long enough time scale to affect more than one spike. This requires merely that the noise have a lower frequency spectrum than the visual signal after temporal filtering by neural mechanisms, as in simulation C above. In either case, observing extended patterns of spikes would reveal that response patterns were more predictable than if the noise had been independent from bin to bin.

Although some cells encoded information in temporal patterns, we found other cells with a net temporal redundancy ($Z < 0$). The largest such effect, $Z = -9\%$, was from a cell with a high burst fraction (15%). LGN neurons fire distinct bursts (Sherman, 1996), which have been shown to carry visual information efficiently (Reinagel et al., 1999). Because bursts have a stereotyped structure, they are internally redundant. Any spike in the burst conveys much the same information as any other. Thus, when we assume temporal independence ($L = 1$), we overestimate the total information in the burst (see simulation B above). We therefore suggest that these bursts contribute a negative term to the pattern correction Z . This idea is consistent with our finding of a significant negative correlation between burst fraction and Z (Fig. 5*c*). Nonetheless, bursting

and temporal redundancy do not preclude high rates of information; the highest information rate in our sample (98 bits/sec at $\delta\tau = 1$ msec) was from a cell with a high burst rate (14%) and a negative pattern contribution ($Z = -2.5\%$).

For all cells in this study, the majority of the encoded information was contained in the temporal precision of firing alone; whether negative or positive, the pattern correction Z was never greater than 25%. Nonetheless, short-range temporal interactions served to increase visual information substantially for many cells (Fig. 5). This indicates that, in those cells, any redundancy between time bins was more than compensated by synergistic effects. We have no reason to think that this result is particular to LGN neurons. Instead, it may be found that neurons in many parts of the brain can exploit both precise spike timing and temporal patterns of spikes to encode useful information. This suggests a possible function for the known sensitivity of neurons to the precise timing and patterns of spikes in their input (Usrey et al., 2000).

Appendix

We define a pattern correction term Z as the difference between the extrapolated information rate for infinitely long words, $I(\lim L \rightarrow \infty)$, and the information rate in the average individual time bin, $I(L = 1)$. This correction depends on one parameter, the temporal resolution $\delta\tau$

$$Z(\delta\tau) \equiv \lim_{L \rightarrow \infty} I(L, \delta\tau) - I(L = 1, \delta\tau) \quad (3)$$

The term $I(L = 1, \delta\tau)$ represents the estimate of information rate that would be obtained on the approximation that time bins are statistically independent within the spike train. The term $I(\lim L \rightarrow \infty, \delta\tau)$ represents the true information rate when all such statistical dependence is taken into account. Thus, Z is a measure of how the relationships between time bins (temporal patterns in the spike train) affect the encoding of stimulus information.

For the interested reader, we derive below an alternative, equivalent expression for Z directly in terms of the statistical structure within spike trains.

The need for Z arises because of the statistical dependence between time bins within the spike train. This property of a spike train can be measured by the average mutual information between a single bin and all other time bins. We will call this I_{int} , for “internal” information. I_{int} measures the average reduction in uncertainty about one bin if you know the others, which is to say, the entropy of a single bin minus the entropy of a bin given all other bins. The average entropy of one bin given all others is equal to the average entropy per bin in the entire sequence (Shannon, 1948, his Theorem 5). Thus, we can write I_{int} in the terms of Eq. 2 as follows:

$$I_{\text{int}}(\delta\tau) = H(L = 1, \delta\tau) - \lim_{L \rightarrow \infty} H(L, \delta\tau) \quad (4)$$

We stress that I_{int} is not information about the stimulus. Rather, it measures the predictability of one time bin of the response from the rest of the response. Note that I_{int} takes into account statistical dependencies of all orders, not just those revealed by an autocorrelation function. In the following discussion, the dependence of all values on the choice of $\delta\tau$ is assumed but not indicated.

If time bins are statistically independent of one another, then $H(L = 1) = H(\lim L \rightarrow \infty)$ and therefore $I_{\text{int}} = 0$. Otherwise, $H(L = 1)$ is the maximum entropy, consistent with the observed probability distribution of spike counts in individual bins, and therefore sets an upper bound on the true entropy: $H(L = 1) \geq H(\lim L \rightarrow \infty)$. This inequality guarantees that $I_{\text{int}} \geq 0$.

A simple rearrangement of Equation 4 produces

$$\lim_{L \rightarrow \infty} H(L) = H(L = 1) - I_{\text{int}} \quad (4')$$

By substituting the right side of Equation 4' for both H_{total} and

H_{noise} in Equation 1, we can rewrite the mutual information between the response and the stimulus as follows:

$$\begin{aligned} \lim_{L \rightarrow \infty} I(L) &= \lim_{L \rightarrow \infty} H_{\text{total}} - \lim_{L \rightarrow \infty} H_{\text{noise}} \\ &= [H_{\text{total}}(L = 1) - I_{\text{int, total}}] - \\ &\quad [H_{\text{noise}}(L = 1) - I_{\text{int, noise}}] \\ &= [H_{\text{total}}(L = 1) - H_{\text{noise}}(L = 1)] + \\ &\quad [I_{\text{int, noise}} - I_{\text{int, total}}] \\ &= I(L = 1) + [I_{\text{int, noise}} - I_{\text{int, total}}] \end{aligned}$$

$$\lim_{L \rightarrow \infty} I(L) - I(L = 1) = [I_{\text{int, noise}} - I_{\text{int, total}}]$$

Finally, substituting Z from Equation 3, we arrive at an alternative expression for the pattern correction Z in terms of two measurements of I_{int} :

$$Z = I_{\text{int, noise}} - I_{\text{int, total}} \quad (5)$$

This formalizes the intuitive notion that the pattern correction Z is entirely attributable to the statistical dependence between bins within spike trains. Recall that all values in Equation 5 depend on the temporal resolution $\delta\tau$, not indicated.

Although both $I_{\text{int, noise}}$ and $I_{\text{int, total}}$ must be positive, the difference between them, Z , may be positive or negative. This is why $I(L = 1)$ is neither an upper bound nor a lower bound on the true information rate. To the extent that the dependence between bins renders the visual information in different bins redundant, $L = 1$ will produce an overestimate, our estimate will decrease with increasing L , and the pattern correction will be negative, $Z < 0$. On the other hand, to the extent that relationships between bins are used to encode stimulus information, $L = 1$ will produce an underestimate, and our estimate will increase with increasing L , reflecting a positive pattern correction, $Z > 0$.

In the special case of a spike train in which the time bins are statistically independent, $I_{\text{int, noise}} = 0$ and $I_{\text{int, total}} = 0$, leading trivially to $Z = 0$ (Eq. 5) and thus $I(\lim L \rightarrow \infty) = I(L = 1)$. However, it is also possible to observe $Z = 0$ even when there is statistical dependence between bins (whenever $I_{\text{int, noise}} = I_{\text{int, total}}$). Temporal structure present in the spike trains may contain a mixture of temporal patterns that encode information synergistically and others that encode stimuli redundantly.

Note added in proof

Two different variants of the direct-information method have been used recently to support similar conclusions for visual neurons in the fly (Brenner et al., 1999) and visual cortical neurons in the primate (Reich et al., 2000).

REFERENCES

- Abeles M, Prut Y, Bergman H, Vaadia E (1994) Synchronization in neuronal transmission and its importance for information processing. *Prog Brain Res* 102:395–404.
- Adrian ED, Zotterman Y (1926) The impulses produced by sensory nerve endings. II. The response of a single nerve end-organ. *J Physiol (Lond)* 61:151–171.
- Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1:371–394.
- 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.
- Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. *Science* 252:1854–1857.
- Brenner N, Strong SP, Koberle R, Bialek W, de Ruyter van Steveninck R (1999) Symbols and synergy in a neural code. Available on the xxx.lanl.gov archives, physics/9902067.
- Buračas GT, Zador AM, DeWeese MR, Albright TD (1998) Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron* 20:959–969.
- Cattaneo A, Maffei L, Morrone C (1981) Two firing patterns in the

- discharge of complex cells encoding different attributes of the visual stimulus. *Exp Brain Res* 43:115-118.
- Dan Y, Alonso JM, Usrey WM, Reid RC (1998) Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat Neurosci* 1:501-507.
- de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W (1997) Reproducibility and variability in neural spike trains. *Science* 275:1805-1808.
- DeBusk BC, DeBruyn EJ, Snider RK, Kabara JF, Bonds AB (1997) Stimulus-dependent modulation of spike burst length in cat striate cortex cells. *J Neurophysiol* 78:199-213.
- DeWeese M (1996) Optimization principles for the neural code. In: *Advances in neural information processing systems 8* (Hasselmo M, ed), pp 281-287. Cambridge, MA: MIT.
- Dimitrov A, Miller J (2000) Natural time scales for neural coding. In: *Computational neuroscience: trends in research 2000* (Bower JM, ed) Amsterdam: Elsevier.
- Eckhorn R, Pöpel B (1975) Rigorous and extended application of information theory to the afferent visual system of the cat. II. Experimental results. *Biol Cybern* 17:71-77.
- Ferster D, Spruston N (1995) Cracking the neuronal code. *Science* 270:756-757.
- Funke K, Wörgötter F (1997) On the significance of temporally structured activity in the dorsal lateral geniculate nucleus (LGN). *Prog Neurobiol* 53:67-119.
- Heller J, Hertz J, Kjaer T, Richmond B (1995) Information-flow and temporal coding in primate pattern vision. *J Comput Neurosci* 2:175-193.
- Levine MW, Zimmerman RP, Carrion-Carriere V (1988) Variability in responses of retinal ganglion cells. *J Opt Soc Am A* 5:593-597.
- Lu SM, Guido V, 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.
- MacKay D, McCulloch WS (1952) The limiting information capacity of a neuronal link. *Bull Math Biophys* 14:127-135.
- McClurkin JW, Gawne TJ, Optican LM, Richmond BJ (1991) Lateral geniculate neurons in behaving primates. II. Encoding of visual information in the temporal shape of the response. *J Neurophysiol* 66:794-808.
- Meister M, Lagnado L, Baylor D (1995) Concerted signaling by retinal ganglion cells. *Science* 270:1207-1210.
- Reich D, Victor J, Knight B (1998) The power ratio and the interval map: spiking models and extracellular recordings. *J Neurosci* 18:10090-10104.
- 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.
- Reich DS, Mechler F, Purpura KP, Victor JD (2000) Interspike intervals, receptive fields, and information encoding in primary visual cortex. *J Neurosci* 20:1964-1974.
- Reid RC, Victor JD, Shapley RM (1997) The use of m-sequences in the analysis of visual neurons: linear receptive field properties. *Vis Neurosci* 14:1015-1027.
- Reinagel P, Godwin D, Sherman SM, Koch C (1999) Encoding of visual information by LGN bursts. *J Neurophysiol* 81:2558-2569.
- Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W (1997). *Spikes: exploring the neural code*. Cambridge, MA: MIT.
- Sejnowski TJ (1995) Pattern recognition—time for a new neural code. *Nature* 376:21-22.
- Shadlen M, Newsome W (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18:3870-3896.
- Shadlen MN, Newsome W (1994) Noise, neural codes and cortical organization. *Curr Opin Neurobiol* 4:569-579.
- Shadlen MN, Newsome WT (1995) Is there a signal in the noise? *Curr Opin Neurobiol* 5:248-250.
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379-423, 623-656.
- Sherman SM (1996) Dual response modes in lateral geniculate neurons: mechanisms and functions. *Vis Neurosci* 13:205-213.
- Stevens CF, Zador A (1995) Neural coding: the enigma of the brain. *Curr Biol* 5:1370-1371.
- Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W (1998) Entropy and information in neural spike trains. *Phys Rev Lett* 80:197-200.
- Theunissen F, Miller JP (1995) Temporal encoding in nervous systems—a rigorous definition. *J Comput Neurosci* 2:149-162.
- Usrey WM, Alonso JM, Reid RC (2000) Synaptic interactions between thalamic inputs to simple cells in cat visual cortex. *J Neurosci* 20:5461-5467.
- 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 (1999) Temporal aspects of neural coding in the retina and lateral geniculate. *Network* 10:R1-R66.
- Victor JD, Purpura K (1996) Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J Neurophysiol* 76:1310-1326.
- Victor JD, Purpura K (1997) Metric-space analysis of spike trains—theory, algorithms, and application. *Network* 8:127-164.
- Warzecha AK, Kretzberg J, Egelhaaf M (1998) Temporal precision of the encoding of motion information by visual interneurons. *Curr Biol* 8:359-368.
- Zador A (1996) Information through a spiking neuron. In: *Advances in neural information processing systems 8* (Hasselmo M, ed), pp 75-81. Cambridge, MA: MIT.