Efficient Discrimination of Temporal Patterns by Motion-Sensitive Neurons in Primate Visual Cortex

Giedrius T. Buračas,*† Anthony M. Zador,*
Michael R. DeWeese, and Thomas D. Albright*
*Howard Hughes Medical Institute
and Sloan Center for Theoretical Neurobiology
The Salk Institute for Biological Studies
La Jolla, California 92037

Summary

Although motion-sensitive neurons in macaque middle temporal (MT) area are conventionally characterized using stimuli whose velocity remains constant for 1–3 s, many ecologically relevant stimuli change on a shorter time scale (30–300 ms). We compared neuronal responses to conventional (constant-velocity) and time-varying stimuli in alert primates. The responses to both stimulus ensembles were well described as rate-modulated Poisson processes but with very high precision (∼3 ms) modulation functions underlying the time-varying responses. Information-theoretic analysis revealed that the responses encoded only ∼1 bit/s about constant-velocity stimuli but up to 29 bits/s about the time-varying stimuli. Analysis of local field potentials revealed that part of the residual response variability arose from "noise" sources extrinsic to the neuron. Our results demonstrate that extrastriate neurons in alert primates can encode the fine temporal structure of visual stimuli.

Introduction

The most popular experimental approach (Adrian, 1926; Hubel and Wiesel, 1962) to studying the neuronal encoding of sensory stimuli makes use of stimuli whose properties are constant over one to several seconds. In this paradigm, response properties are summarized by a tuning curve, which presents the firing rate as a function of some stimulus parameter. The "firing rate" is typically defined as the number of spikes (averaged over repeated stimulus presentations) generated in a time window of several seconds, and the optimal stimulus is then defined in terms of the stimulus features that elicit the highest firing rate.

Under ecological conditions, visual stimuli rarely remain constant for longer than a fraction of a second: eye movements, self-motion, and object motion render the visual scene highly variable. Moreover, an organism outside of the laboratory often does not have the luxury of averaging several stimulus presentations over extended time intervals, but must instead base its (sometimes life-or-death) decisions on the neuronal response during a single "trial" of 30–300 ms duration. Under such conditions, the fidelity of neuronal responses becomes a critical factor. Because classical tuning curves depend only on the mean firing rate, and not on the variability about that mean, they provide no insight into neuronal

fidelity. Information-theoretic measures, by contrast, express not only what stimulus characteristics are encoded in a spike train but also how well they are encoded (see, e.g., Rieke et al., 1997).

We have explored how motion-sensitive neurons in the middle temporal (MT) cortex of alert monkeys encode rapidly and unpredictably changing stimuli. We have adopted an approach in which encoding is defined in terms of the stimulus features that a neuron is best able to discriminate rather than in terms of the features that drive it to fire fastest. We quantitate an average neuronal discriminative power in terms of the information conveyed by the spike train about the stimulus. In this approach, the optimal stimulus is defined as the one that produces the highest information rate. We therefore searched for the stimulus ensemble that maximized this rate.

Our major finding is that neurons in area MT of alert monkeys encoded information at much higher rates when driven by stimuli with a rich temporal structure than when driven by conventional (constant-motion) stimuli (Bair et al., 1996, 1997). The increased information was due to the enhanced stimulus-locked timing precision of the neuronal response; trial-to-trial variability in the spike count was the same for both types of stimulus. Analysis of local field potentials suggested that at least part of the spike-count variability was due to trial-to-trial variability in the inputs driving the neurons rather than noise sources intrinsic to the neuron under study.

We used two methods to calculate information rates. The direct method estimates the total information available in the spike train about all aspects of the stimulus, while the reconstruction method estimates the information about some specific stimulus parameter(s) (e.g., direction of motion). A comparison of these rates suggested that about half of the available information encoded the direction of stimulus motion under our conditions.

Portions of this work have previously appeared in conference abstracts (Buračas et al., 1996, 1997, Soc. Neurosci., abstracts).

Results

We present our results as a contrast between neuronal encoding of constant-motion and variable-motion stimuli by area MT neurons of alert monkeys. The analysis of the encoding of the constant-motion stimulus serves as a baseline against which the results from the variable-motion stimulus are compared.

Encoding of Constant-Motion Stimuli in Area MT Responses

Neurons in area MT of primate visual cortex encode information about the velocity (direction and speed) of moving images (Maunsell and Van Essen, 1983; Albright, 1984). We recorded extracellularly from individual neurons in area MT of rhesus monkeys (*Macaca mulatta*)

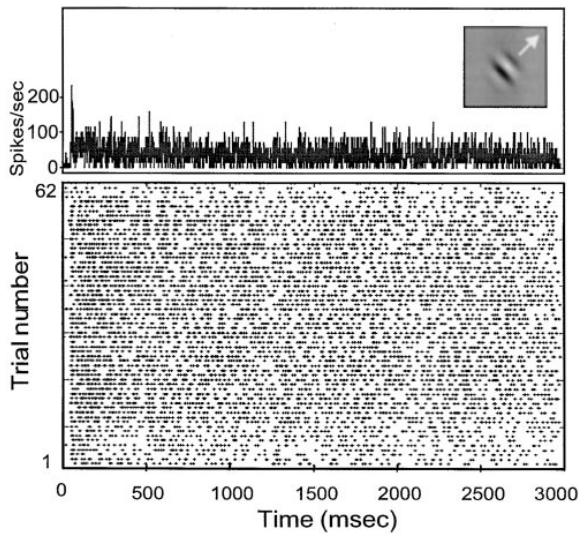†To whom correspondence should be addressed.

Figure 1. MT Neurons Produce Highly Variable Responses When Presented with Constant Motion Stimuli

The top curve shows a typical poststimulus time histogram (PSTH) obtained using bins of 1 ms width for 62 presentations of the stimulus, the spatial profile of which is illustrated in the inset. Rasters illustrating spike occurrences on individual trials are shown in the bottom panel.

who viewed moving sinusoidal gratings windowed by a stationary Gaussian envelope (Gabor patch; Figure 1, inset). We will refer to this as the "constant-motion stimulus." As expected, these neurons responded selectively to different directions of stimulus motion (Figure 2a).

The efficiency with which neurons encode the direction of stimulus motion can be quantified by the relative entropy $I(\theta)$ (see, e.g., Cover and Thomas, 1991). Intuitively, $I(\theta)$ can be interpreted as the "information per direction" (see Experimental Procedures for details). Figure 2b shows the information per direction $I(\theta)$, which has units of bits, for the first second following stimulus presentation. $I(\theta t)$ spans only a 2-fold range, in contrast with the nearly 10-fold range of the firing rate (Figure 2a). Thus, the informational tuning curve indicates that the spike train carries some information about all directions.

The number of spikes fired during a 1 s trial is insufficient to discriminate reliably among the eight possible stimulus directions. The discriminative capacity is given by the total information rate, which is obtained by averaging the information per direction over all eight stimulus directions. Perfect discrimination between all eight directions would require $\log_2 8 = 3$ bits, while this MT neuron conveys only 1.38 bits of information about the direction of stimulus motion in 1 s. Similar results were obtained from other area MT neurons (n = 12, mean ± SD = 0.89 ± 0.29 bits/s; Figure 2C). One bit is sufficient to discriminate between $2^{\#bits} = 2^1 = 2$ possibilities with 100% accuracy or among a greater number of possibilities with lower accuracy (see Experimental Procedures, Signal Detection versus Information Theory). Since the firing rate (averaged over all eight directions) is about 40 spikes/s, the average information transmitted is 0.025
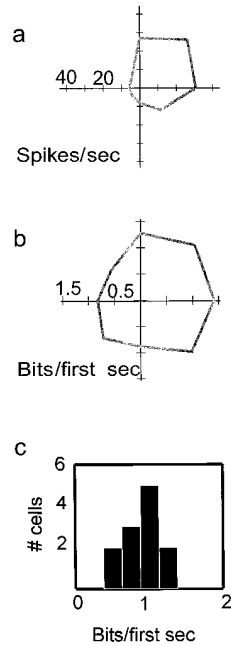

Figure 2. The Information Transmitted by a Neuron Is Small in Response to Constant Motion Stimuli

(a) Neurons in area MT are directionally sensitive. A typical direction tuning curve derived from the average firing rate during presentation of the constant-motion stimulus is shown in polar coordinates. The angular axis represents the direction of stimulus motion; the radial axis represents the average firing rate (spikes/s). Note that this neuron is almost 10 times more responsive to stimuli moving in the preferred direction than the antipreferred direction.

(b) An MT neuron can be informative about directions that are not near the peak of its conventional directional tuning curve. In this polar plot, the radial axis represents $I(\theta)$, which can be thought of as the "information per direction" conveyed by the spike count in the first second after stimulus presentation (bits per first second) (see Experimental Procedures). Note that the transmitted information is within a factor of two of its greatest value for all directions, despite nearly an order of magnitude in range of average firing rates illustrated in (a).

(c) The spike count does not convey much information about the direction of constant motion stimuli, even after a full second following the stimulus onset. The bar graph is a frequency distribution across cells of information about direction of motion after 1 s of stimulus onset (mean = 0.89 bits per first second; n = 12).

bits/spike. These values are in agreement with estimates of the information transmitted by cells in the retina (Fitz-Hugh, 1957) and inferior temporal cortex (Optican and Richmond, 1987) in response to constant stimuli.

The discriminative capacity of a code based on spike count is limited by two factors. First, the dynamic range of the spike rate is limited, so that the maximum number of distinct directions that could possibly be distinguished based on the spike count in any fixed time window is bounded by the maximum number of spikes the neuron can fire in that window. For example, if a neuron can sustain a rate no higher than 100 spikes/s, then on a given trial the spike count in 1 s must be one of 101 possible values (0, 1, ..., 100). Therefore, the spike count can in principle distinguish at most among 101 directions corresponding to $\log_2 101 \approx 6.7$ bits/s. Second, the spike train is noisy; i.e., it exhibits considerable trial-to-trial variability, which is consistent with a Poisson
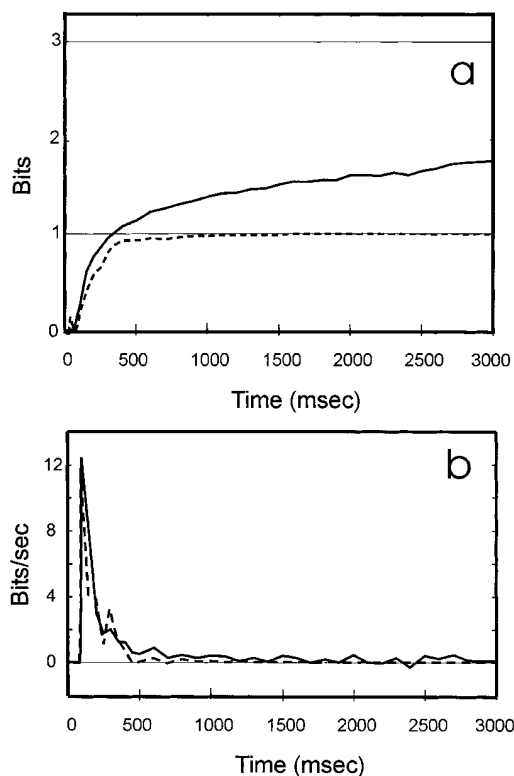
Figure 3. Most of the Information about the Direction of the Constant Motion Stimulus Is Available Immediately after the Stimulus Onset

(a) The information for the constant motion stimulus grows rapidly for only about the first 250 ms. The curves represent the accumulated information conveyed by the spike count about stimulus direction as a function of the time since stimulus onset for two different stimulus ensembles. The dashed curve shows that the information rapidly saturates when the stimuli could have moved in one of two equally likely directions. The total amount of information about direction available in this stimulus ensemble was $\log_2(2) = 1$ bit (lower thin horizontal line). The solid curve represents the same quantity for a stimulus that moved in one of eight equally likely directions; the available information in this ensemble was $\log_2(8) = 3$ bits (upper thin horizontal line). Even after the full 3 s of stimulation, the spike count contains just over half of the available information. Note that a 3-fold increase in the available information results in a small difference between these curves.

(b) The information rate for the constant motion stimulus is high at first but quickly decays. The dashed curve shows the temporal derivative of the accumulated information (i.e., the instantaneous information rate) corresponding to the dashed curve in (a) (two-direction ensemble). The solid curve shows the information rate corresponding to the solid curve in (a) (eight-direction ensemble).

process (see Figure 1 and below). In this case, the amount of information drops to at most 2.2 bits/s, corresponding to four to five directions. In practice, the information per direction is even lower because direction is not optimally mapped to a neuron's firing rate.

Figure 3a shows that the information is a rapidly saturating function of the time after stimulus onset; i.e., the slope of the curve is steepest immediately after the stimulus onset (Figure 3b). Although the total information transmitted by the neuron of Figures 1–3 in 1 s was only 1.38 bits, over 50% of that information (0.76 bits) was available within the first 200 ms (Figure 3a, solid line).

Furthermore, when the stimulus is constrained to move in one of two directions (along the preferred/antipreferred axis), an even larger fraction (80%) of the total information was available by 200 ms (Figure 3a, dashed line). Thus, particularly for short intervals, the information available about the two- and eight-direction stimuli differs only slightly. Similar saturation has been reported in earlier studies of the primate inferior temporal cortex (Optican and Richmond, 1987; Tovee et al., 1993).

Figure 3b shows the instantaneous information rate for both stimulus conditions obtained by taking the temporal derivative of the information curves in Figure 3a. The instantaneous information rate rapidly attains a peak value of 12 bits/s but quickly declines to a steady state of nearly 0 bits/s, so that the accumulated information is small at the end of the trial. The early peak in the instantaneous information rate raised the question of whether the neuron could sustain this rate for extended periods of time.

Encoding of Rapidly Changing Stimuli

The low information rate ($\sim$1 bit/s, 0.025 bits/spike) elicited by the constant-motion stimulus might result from a mismatch between the impoverished stimulus set we used and the true discriminative capacity of MT neurons. Such a mismatch has been reported in the blowfly, where the motion-sensitive H1 neuron transmits information at very high rates in response to stimuli with rich temporal structure but not in response to more conventional constant-motion stimuli (de Ruyter van Steveninck et al., 1997). That such a mismatch also accounts for the low information rates seen in MT neurons is not at all self-evident: flies have only a handful of motion-sensitive neurons, so it would not be surprising if they employed a computational strategy very different from that used by monkeys, a strategy in which nearly every spike from every neuron was significant. However, the steep initial increase of the information (Figure 3) suggested that constant-motion stimuli were similarly mismatched to the primate visual system.

In order to test whether the transiently elevated information rate seen in the first few hundred milliseconds could be sustained over longer periods, we designed a set of richer visual stimuli. These stimuli switch randomly between a neuron's preferred and antipreferred directions on an ecologically relevant time scale (about every 30–300 ms; Figure 4a). We constrained the range of directions to only two values since, as noted above, an average MT neuron is unable to discriminate reliably among more than two directions within this short time interval.

To quantitate the capacity of MT neurons to discriminate time-varying stimuli, we first applied a direct information estimation method (DeWeese, 1996; Stevens and Zador, 1996; de Ruyter van Steveninck et al., 1997; see Figure 5a) to 28 cells recorded from two alert monkeys. The direct method compares response reliability to response repertoire. Response reliability was assessed by presenting the same time-varying stimulus repeatedly on many trials (Figure 4b), while repertoire was assessed by presenting stimuli with different temporal structure on different trials (Figure 4c). Intuitively,
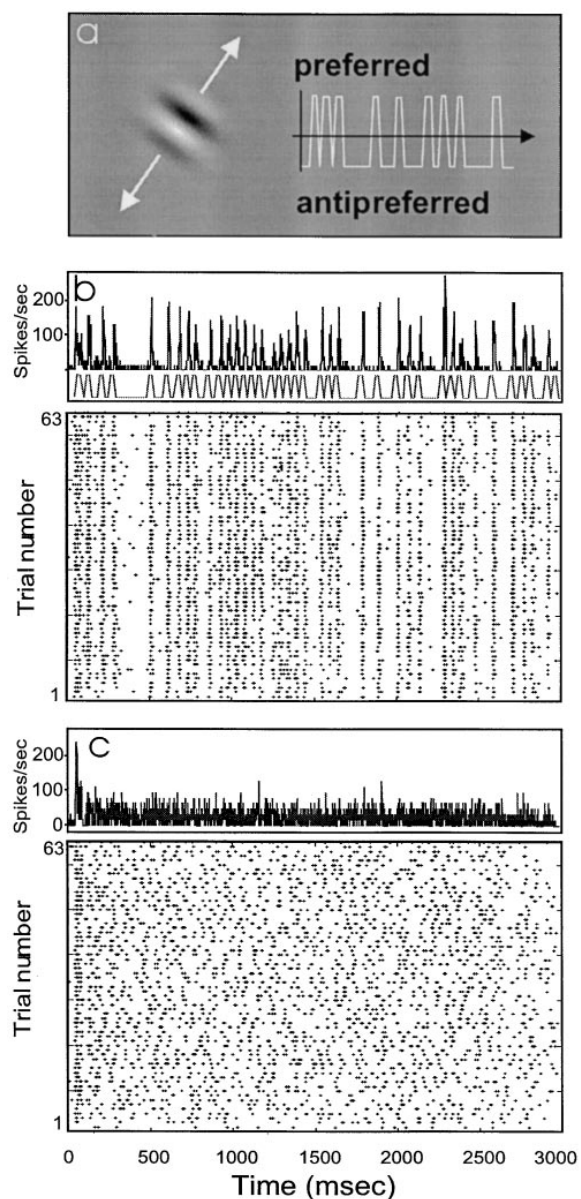
Figure 4. MT Neurons Can Discriminate Time-Varying Stimuli Much More Effectively Than Conventional Constant-Motion Stimuli

(a) The stimulus switched rapidly and stochastically between the preferred and antipreferred directions (see Experimental Procedures).

(b) Repeated presentations of the same time-varying stimulus produced stimulus-locked responses with high timing precision. The top curve shows the PSTH resulting from the time-varying stimulus depicted by the lower curve; response rasters are displayed below. These responses were used to estimate the conditional entropy (see Experimental Procedures). The stimulus trace was aligned with the response peaks by shifting it by the latency of 60 ms. The stimulus parameters were a = 2, b = 1, p = 0.33.

(c) Presentation of different instantiations of the stimulus, drawn randomly from the same ensemble as in (b), produced highly variable responses. The PSTH (top) is shown above the response rasters. These responses were used to estimate the total entropy.

the information conveyed by the spike train is large if (1) the response repertoire is rich and (2) the response reliability is high. The repertoire can be quantified in

terms of the total entropy of the spike train, the reliability in terms of its conditional entropy, and the Shannon mutual information as the difference between these entropies (see Experimental Procedures).

The results shown in Figure 4 demonstrate that MT neurons can sustain high information rates if the stimulus has a rich temporal structure. The responses illustrated in Figure 4, for example, convey 1.4 bits/spike, which at a spike rate of 21 spikes/s yields 29 bits/s. Similar measures were computed for each of 28 neurons and are plotted on the right side of Figure 6a. The mean of this distribution is 1.5 bits/spike (12 bits/s; Figure 6b), which implies that on average an MT neuron can discriminate about 2^{12} ($= 4 \times 10^4$) different temporal patterns of 1 s duration. This remarkable discrimination performance is exhibited only when the stimulus offers novelty at every instant in time, i.e., when the total stimulus entropy grows linearly with time. It is interesting to note that the information rate we observed is comparable to the 1–3 bits/spike reported in a wide variety of other systems (see, e.g., Rieke et al., 1997) but is much higher than early reports for neurons in monkey cortex (<0.1 bits/spike; Optican and Richmond, 1987; cf. Bair et al., 1997).

The direct method permits information to be estimated without specifying which aspect of the stimulus is being encoded in the spike train. In order to measure the amount of information conveyed specifically about the stimulus direction at each frame, we also employed a reconstruction method (de Ruyter van Steveninck and Bialek, 1988; Rieke et al., 1997; see Figure 5b). This method gives lower information estimates (left side of Figure 6a) than the direct method for two reasons. First, it reflects only information associated with a single stimulus parameter (direction at each video frame). This method discards information that a neuron can carry about other stimulus parameters, such as the spatial phase of the Gabor patch and the temporal phase of monitor refresh. Second, the estimate of the information rate also depends on the quality of the reconstruction algorithm—to the extent that the reconstructor is suboptimal, this method underestimates the information rate.

The average information obtained from 26 cells using this reconstruction method is only 0.6 bits/spike (5.5 bits/s; Figures 6b and 6c, center). Therefore, when we relax our assumptions about what features are encoded and how they are encoded by using the direct method, we obtain twice the information rate found with the reconstruction method (cf. Figures 6b and 6c, center and right).

Analysis of Response Variability

The trial-to-trial variability of the spike count, and the spike timing precision, are the two critical factors that together determine a neuron's discriminative power. The observed spike-count variability, as quantified by the Fano factor (the variance divided by the mean of the spike count) only rarely fell below that predicted by a Poisson process, which has a Fano factor of unity. The average Fano factors were 1.27, 1.23, and 1.36 for constant-motion, repeated, and random stimuli, respectively, and ranged from 0.66–2.36 using a time window

a Direct Method

$$I(s,\{t\})=H(\{t\})-H(\{t\}|s)$$

b Reconstruction Method
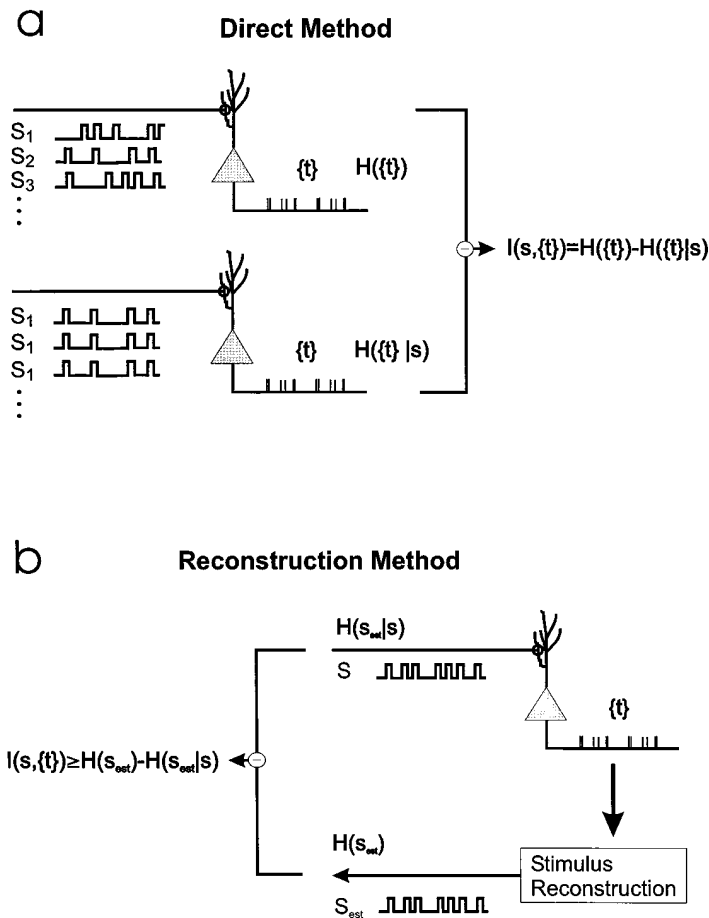

$$I(s,\{t\})\geq H(s_{est})-H(s_{est}|s)$$

Figure 5. Two Methods Used to Measure Information

(a) Direct information estimation methods do not require the experimenter to know what aspects of the stimulus are encoded, nor how they are encoded in the spike train. Our information estimate is the difference between the entropy calculated from responses to randomly generated stimuli [$H(\{t\})$] and the entropy of responses to a specific repeated stimulus drawn from the same distribution [$H(\{t\}|s)$] (see Experimental Procedures).

(b) Reconstruction methods for estimating information can provide lower bounds to the information and are useful for determining about which aspects of the stimulus the cell is informative. Our reconstruction method estimates information as the difference between the entropy of an optimized nonlinear reconstruction of the random stimulus [$H(s_{est})$] and the entropy of the errors made by the reconstruction [$H(s_{est}|s)$].

of 100 ms. The Fano factor showed a consistent increase with window size (Figure 7a; see also Teich et al., 1990). The Fano factor was also somewhat lower (mean ± SD = 1.1 ± 0.3) in the 40 ms windows centered on the stimulus-locked peaks of the response (see Figure 7b) but still near that expected from a Poisson process. Thus, individual MT neurons remain highly variable even for the repeated time-varying stimulus. By contrast, vertebrate retinal ganglion neurons (Berry et al., 1997) and motion-sensitive neurons in the fly (de Ruyter van Steveninck et al., 1997) can exhibit low, sub-Poisson spike-count variability. Since the Fano factors we observed were about the same for the constant and rapidly varying stimulus ensembles, they could not have been responsible for the dramatic increase in the information rate observed in the latter case.

By contrast, the timing precision of stimulus-locked responses was very high, in some cases as high as 1.9 ms (Figure 7c, black bars). Focusing solely on the first spike yields slightly higher timing precision (Figure 7c, gray bars; see also Bair and Koch, 1996). In fact, the spike timing precision for most neurons was several times smaller than our stimulus timing interval (16.7 ms, limited by the 60 Hz video refresh rate of our monitor), suggesting that the estimated information rate might have been even higher if stimuli were modulated on a finer temporal scale. Thus, it was the high temporal precision, and not the spike-count variability, that accounted for the high information rates observed in the

responses to the rapidly varying ensemble (see also Reich et al., 1997).

Sources of Response Variability

The observed trial-to-trial variability in spike count (Figure 4) must arise from some combination of two sources. It might reflect some source of "intraneuronal noise," i.e., variability specific to the neuron whose activity is being monitored. In view of the high fidelity of spike transduction (Bryant and Segundo, 1976; Mainen and Sejnowski, 1995), the dominant mechanism underlying intraneuronal noise is probably synaptic unreliability (Allen and Stevens, 1994; Stratford et al., 1996; Dobrunz and Stevens, 1997; Zador, 1998). Alternatively, the variability could reflect "extraneuronal noise," i.e., trial-to-trial variability already present among the inputs to the neuron whose activity is being monitored. Plausible sources of extraneuronal noise include intrinsic variability in the output of the neurons (e.g., in area V1) driving those in MT, the inconstancy of the retinal stimulus resulting from eye movements (Gur et al., 1997), changes in the state of attention or arousal, and uncontrolled activity outside the classical receptive field.

In order to distinguish these two broad classes of noise, we recorded local field potentials (LFPs) and single unit activity simultaneously from a single electrode. The LFP reflects synaptic activity in the vicinity of the recording electrode (Gray, 1994; see Experimental Procedures) and thereby serves as a gross measure of the
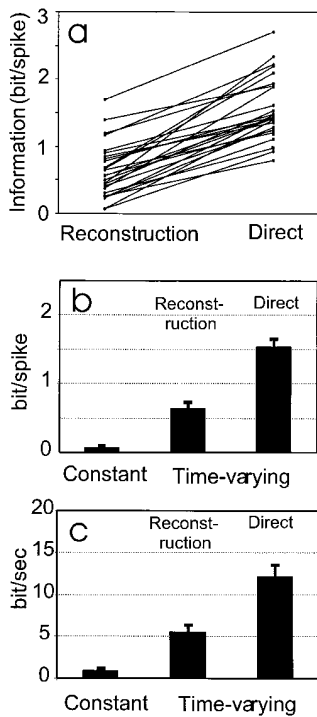
Figure 6. MT Neurons Can Transmit High Rates of Information about Temporally Rich Stimuli

(a) The plot compares the information (bits/spike) obtained with the reconstruction (left) and direct (right) methods. Lines connect pairs of values for the same cell.

(b) Individual spikes convey more information about the timing of changes in direction of motion of time-varying stimuli than about the exact direction of constant-motion stimuli. The left bar represents the average information conveyed by one spike (bits/spike) during 1 s of stimulation by the constant-motion stimulus. The middle and right bars represent the information per spike for time-varying stimuli using the reconstruction method and direct method, respectively. Error bars represent standard errors across cells.

(c) Time-varying stimuli elicit higher information rates than do constant-motion stimuli. The bars represent information per unit time (bits/s) rather than information per spike as in (b).
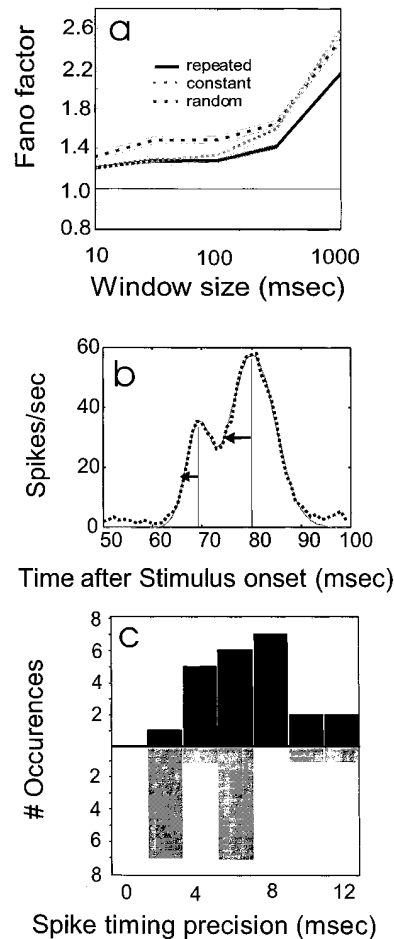

Figure 7. Response Timing Precision Was High Despite the High Trial-to-Trial Variability in the Spike Count

(a) Trial-to-trial spike-count variability was high as indicated by the Fano factor. The Fano factor is defined as the ratio of the variance to the mean of the spike count for some fixed time window. The Fano factor was computed as a function of window size for sliding windows of different lengths for all three stimulus ensembles: constant (gray dotted), random time-varying (black dotted), and repeated time-varying (black solid). Each curve is the mean ($\pm$ SEM) over the 17 neurons for which all three stimulus ensembles were presented. All three curves are above 1 over the full range of time intervals between 10 and 1000 ms. For a Poisson process, the Fano factor is unity for all window sizes (thin horizontal line). The fact that the Fano factor was near or above unity for all window sizes, even for the repeated stimulus ensemble, demonstrates that the spike-count variability is at least as high as that of a Poisson process.

(b) Responses showed high timing precision. The dotted line shows the average response triggered by stimulus motion reversals from the antipreferred to the preferred direction for a typical neuron. The solid line shows a sum-of-two-Gaussians fit. There are two peaks because the preferred motions occurred in pairs in the optimal stimulus for this cell. The arrows represent standard deviations, which quantitate timing precision.

(c) The timing precision was high across all cells. The black bars show the frequency distribution of timing precisions as defined by the standard deviations shown in (b); by this measure, the median timing precision was 6.17 (median-to-quartile differences were −1.49, +2.27 ms, n = 17; this includes standard deviations from both Gaussians as shown in [a]). The gray bars represent the frequency distribution of the timing precision of the first spike occurring inside the response region, which starts two standard deviations before the peak of the first Gaussian; by this measure, the median timing precision was 5.58 ms (median-to-quartile differences were −3.23, +1.03 ms; n = 17).

magnitude of the input driving a region. Because they represent an average of activity over a region, LFPs do not reflect intraneuronal noise, but they might reflect some sources of extraneuronal noise.

As seen in Figure 7b, each reversal of the stimulus motion from the antipreferred to the preferred direction induces an increase in the firing rate one latency ($\sim$60 ms) later. However, only a fraction of trials was associated with a spike during any such stimulus-locked increase in firing rate. We therefore assessed whether the presence or absence of such a spike on any particular trial could be predicted by the state of the local field potential immediately preceding the stimulus-locked increase.

We defined a measure R to quantitate the correlation between trial-to-trial fluctuations in the spike count and the preceding LFP (see Experimental Procedures). R had the following properties. If the LFP was uncorrelated with trial-to-trial variations in the spike count, then R = 0.5. If the average LFP preceding a stimulus-triggered

event was always higher on trials with no spikes than on trials with spikes, then R = 1. Note that R is defined so that it depends only on trial-to-trial correlations between the spike count and the preceding LFP and not on systematic variations in the LFP or the spike count during a single trial due to the details of the stimulus.

For the cell shown in Figure 7a, R = 0.74. For all nine cells in which LFPs were recorded, R > 0.5 (R = 0.68 ± 0.13 SD); of these, the effect was statistically significant (p < 0.01) for six of nine cells (for this subset, R = 0.75 ± 0.07 SD). These results indicate that at least some of the trial-to-trial variability in the spike count can be accounted for by population activity not locked to the stimulus.

Discussion

Our major finding is that individual neurons in area MT of alert monkeys are better able to discriminate between stimuli with rich temporal structure than constant-motion stimuli that differ only in direction. These neurons encoded stimulus direction rather coarsely (two to three directions, corresponding to 1–1.6 bits of information, in 1 s), and temporal integration over long intervals improved discrimination only marginally. Moreover, most of the information associated with the constant motion stimulus was carried by the transient response immediately following stimulus onset. In order to test whether these neurons could sustain the high information rate seen during the stimulus onset transient, we examined their responses to stimuli with rich temporal structure. These neurons encoded the temporal structure of rapidly changing stimuli with very high temporal precision, when temporal structure was defined on an ecologically relevant timescale (equivalent to perfect discrimination of up to $2^{29} \approx 5 \times 10^8$ equally likely temporal patterns of 1 s duration or imperfect discrimination of a larger number of stimuli). This remarkable discrimination power was due to the high temporal precision of the stimulus-locked response. By contrast, the trial-to-trial variability of the spike count was high and consistent with a rate-modulated Poisson process.

Precision and Variability of Signaling by MT Neurons

The first suggestion that area MT neurons might be sensitive to the fine temporal structure of visual stimuli was provided by Bair and Koch (1996), who found that the response of these neurons to a repeated random-dot stimulus was replicable with high temporal precision (<2 ms) but only when the "stimulus coherence" fell within a certain range. In the current study, we confirm and extend this finding to stimuli with a rich and well-defined temporal structure and suggest that the high timing precision can be attributed to the fine temporal structure of the stimulus. Under these conditions, the temporal precision of responses in area MT is of the same magnitude as that observed in the rabbit retina (Berry et al., 1997). This is a striking fact, considering that individual MT neurons lie a minimum of three synapses beyond retinal ganglion cells.

MT neurons responded with high temporal precision

when challenged with rapidly changing stimuli, but showed high (Poisson) spike-count variability under all conditions (see also Reich et al., 1997). By contrast, the H1 neuron of the fly shows much lower (sub-Poisson) variability when driven by rapidly changing stimuli (de Ruyter van Stevenick et al., 1997). This may reflect differences between the constraints on the motion-processing circuitry of primates and flies. In primate cortex, there are many thousands of motion-sensitive neurons, and they can pool responses to obtain a better estimate of the instantaneous spike rate, whereas flies have only a few and so cannot effectively average. Note that such pooling can readily compensate for high spike-count variability but could not so easily compensate for low temporal precision. Interestingly, the trial-to-trial variability is also low in the vertebrate retina (Berry et al., 1997).

Slow Synaptic Activity in Local Neuronal Population Predicts Spike Occurrence

Previous studies have shown that the local field potential can be used to predict the occurrence of a spike in the visual cortex of anesthetized cats (Arieli et al., 1996). These results suggested that spontaneous activity of single neurons was due to the coherent activation of synaptic input that was not time-locked to the stimulus. The coherent activation was attributed to "ongoing [neuronal] activity," probably due in part to anesthesia-induced synchrony.

The present findings extend those results in two ways. First, our studies were performed in alert animals, so that the coherent activity could not have been due to anesthesia (see also Arieli et al., 1996, Soc. Neurosci., abstract). Second, the stimuli we presented elicited spikes with very high timing precision, so that there were well-defined stimulus-locked periods during which a spike was expected to occur. We found that the LFP was correlated with the presence or absence of a spike during these stimulus-locked periods, suggesting that LFPs possessed stimulus-independent components, which nonetheless could predict spike occurrence.

Our results are also consistent with the observed correlation in the spike count of nearby neurons in area MT of alert monkeys (Zohary et al., 1994). Since the LFP reflects the gross synaptic drive to a region, it would have been similar at two nearby neurons isolated by a single electrode. Since the spike count from each neuron would have been correlated with the LFP, the spike counts of the two neurons would have been correlated with each other. Whether the variability represents "noise" depends on its source, which has not yet been established.

Rate versus Temporal Coding

The nature of the neural code has recently been subjected to renewed scrutiny (see, e.g., Softky and Koch, 1993; Ferster and Spruston, 1995; Shadlen and Newsome, 1995; Stevens and Zador, 1996; Rieke et al., 1997). Part of this controversy arises from ambiguity in the terms "temporal code" and "rate code." According to the classical notion of a rate code, all information in the spike train is available in a single number, the mean

firing rate. This definition was developed to describe responses to stimuli whose properties are constant over one to several seconds; it does not make clear predictions about the responses to stimuli with temporal structure on a time scale short compared to typical interspike intervals.

By contrast, the term temporal code is often taken to refer to any code in which there is information contained in a spike train beyond that available in the mean rate, i.e., any code in which the fine temporal structure of the spike train matters. For example, it has been suggested that single neurons in the primate inferior temporal cortex driven by static stimuli use a form of temporal multiplexing to transmit independent information about different stimulus parameters, such as the shape of components of static images (Optican and Richmond, 1987).

In our experiments, we found that the fine temporal structure of the spike train did encode information but only about the fine temporal structure of the stimulus. Whether this represents an example of temporal coding is essentially a matter of definition. The classical notion of rate coding is well defined under conditions where the stimulus parameter encoded by a neuron changes slowly compared with the typical interspike interval, since under these conditions the mean rate can be estimated from the spike count. The classical notion does not, however, make clear predictions when the stimulus is changing rapidly. We prefer to view our results as consistent with a natural extension of the classical rate coding model to rapidly varying stimuli.

What Do Neurons in MT Encode?

In the usual approach to studying what a neuron encodes, the experimenter studies the effect on firing rate of varying a stimulus parameter. This stimulus parameterization may in some cases not fully characterize the stimulus. By focusing on some aspects of the stimulus to the exclusion of others, this approach implicitly injects the experimenter's preconceived biases about what the neuron encodes.

A striking example of such a partial stimulus parameterization is the popular random dot stimulus, which is usually described by two parameters—direction of apparent motion of random dots and noise level (e.g., Newsome et al., 1989)—but ignores the detailed dynamics of individual dots. The dynamics of individual dots can, however, have a significant and replicable effect on the neuronal response, but these dynamics do not enter into the stimulus description. In effect, some of the information transmitted with high fidelity by MT neurons is discarded as noise.

The dependence on stimulus description can be made explicit by employing information-theoretic methods to dissociate the total available information from the information about specific stimulus parameters. We achieved this dissociation by using two independent methods to estimate the information rate (see Figure 5). The direct method assessed the neuron's overall discriminative power, without reference to the features being discriminated. The direct method is thus independent of any subjective assumptions about which stimulus parameters, given a particular stimulus ensemble, are and are not relevant. The direct method thereby furnishes an objective estimate of the maximal discriminative power a neuron can exhibit on a given stimulus ensemble.

The reconstruction method, by contrast, assesses only the information provided by the neuron's output about some specific stimulus parameter; it discards all information deemed by an experimenter to be irrelevant. A comparison of the two estimates thereby provides a way of testing the validity of our beliefs about the relative importance of various stimulus parameters. In our experiments, the information estimated using the reconstruction method was about half of that based on the direct method (5.5 bits/s versus 12 bits/s). Since the reconstruction method estimates a lower bound that is limited by the quality of the reconstructor (see Experimental Procedures), part of this difference may be due to the possibility that our estimator may have been suboptimal. However, the 2-fold difference in the information estimates suggests that MT neurons may be sensitive not only to the sign of speed/direction reversals at each video frame but also to other dimensions of our stimulus ensemble, such as the spatial phase of the Gabor patch, the monitor refresh phase, and so on.

The lower bound on the information rate (5.5 bits/s) that we obtained using a nonlinear reconstruction method is in close agreement with the 6.7 bits/s reported by Bair et al. (1997) using a different stimulus ensemble and a linear reconstruction method. The two estimates are of the same order despite the fact that the stimuli used by Bair and colleagues had an entropy several times higher than those used in the present study (over 100 bits/s versus 28–60 bits/s). This agreement suggests that the information rate may not depend sensitively on the stimulus ensemble, provided the ensemble is endowed with sufficiently rich temporal structure.

The information-theoretic framework can be viewed as a generalization of the receiver operating characteristics (ROC) analysis already widely used to quantitate stimulus discriminability (Tolhurst et al., 1986; Newsome et al., 1989; see Experimental Procedures). It might be argued that ROC analysis better reflects the behaviorally relevant discriminative capacity than information theoretic measures. This view arises because ROC is the natural analysis for a common experimental paradigm, namely the two-alternative forced-choice paradigm. A forced-choice paradigm may not, however, always be an appropriate model of natural behavior. Information theory may provide a more natural measure for behavioral paradigms in which the task requires processing of the fine spatiotemporal structure of the stimulus or when the task cannot be readily formulated in terms of discrimination.

What does a sensory neuron really encode? The time-varying stimuli we tested are suboptimal for MT neurons in the conventional sense because they elicit fewer spikes than the "optimal" stimulus moving at a constant speed in a neuron's preferred direction. However, if we define encoding in terms of discrimination (as measured by information rates), we find that the optimal stimuli are stimuli that possess rich temporal structure.

Experimental Procedures

Stimuli

The stimulus was a Gabor patch subtending 4° of visual angle presented on a 21 in computer monitor with 60 Hz refresh rate (Figure

1a). Motion was achieved by a spatial phase shift occurring each video frame (net speed of 12°/s). Direction tuning was assessed using stimuli whose phase moved with constant speed and direction for 3 s in one of eight different directions (constant-motion stimulus). For the time-varying stimulus, direction reversals from the antipreferred to the preferred direction were followed by a frames in the preferred direction, which were followed by b frames in the antipreferred direction, after which the stimulus became stochastic, randomly switching to the preferred direction with probability p at each frame. For example, $a = 1$, $b = 0$, $p = 1/2$ corresponds to a random coin flip on each frame, with an entropy of 1 bit/frame. The parameters a (1–3 frames), b (0–3 frames), and p (0.28–0.5) were adjusted online to maximize the information rate for each cell. In our experiments, the stimulus entropy ranged from 14–60 bits/s.

Behavioral Paradigm and Recording Techniques
The monkeys were alert, fixating their eyes and passively viewing during stimulus presentation. Eye position was monitored with the scleral eye-coil method. Surgical procedures, single unit recording technique, and data acquisition system were standard and have been described elsewhere (Albright, 1984; Dobkins and Albright, 1994). LFPs were recorded simultaneously with action potentials using a tungsten microelectrode of 200–500 kΩ impedance at 1 kHz. LFPs were low pass filtered (Nyquist cutoff frequency = 625 Hz) and sampled at 1250 Hz.

Data Analysis
We used three data analysis methods derived from information theory (Shannon, 1948; Cover and Thomas, 1991; Rieke et al., 1997).

Information per Direction
For the conventional constant-motion stimulus (Figure 1), we employed Bayes' formula: the conditional probability of stimulus direction θ given spike count z in a fixed time window is $p(\theta|z) = p(z|\theta) p(\theta)/p(z)$. We assumed that $p(z|\theta)$ was distributed according to a homogeneous Poisson process for each stimulus direction. This approximation was justified, because (1) the interspike intervals were nearly exponentially distributed and (2) the Fano factor was close to 1. The relative entropy (Kullback-Liebler distance between $p[z|\theta]$ and $p[z]$) $I(\theta)$ is defined as

$$I(\theta) = \sum_z p(z|\theta)\log_2[p(z|\theta)/p(z)].$$

The mutual information between the stimulus direction τ and the spike count z is the average of $I(\tau)$ over the stimulus distribution

$$I_{dir} = \sum p(\theta)I(\theta),$$

which suggests the interpretation of $I(\theta)$ as the information per direction. We note, however, that this is not the only quantity whose average is the mutual information (cf. Rieke et al., 1997). We limited our discussion to eight directions since the information did not increase appreciably with increasing number of directions.

Direct Method
For the time-varying stimulus, we used both a direct method and a reconstruction method to estimate the mutual information, I. In the direct method, we used two modes of stimulation. First, we repeatedly presented the same 3 s instantiation of the stimulus. This was used to measure the "conditional" entropy, C, of the spike train. Next, we presented different 3 s instantiations of the stimulus, which were all drawn from the same distribution. This was used to measure the "total" entropy, T, of the spike train. The information was then given by $I = T - C$. To estimate the entropies, we counted the number of spikes in successive bins of size dt and constructed "words" (strings of fixed length in which "0" in a bin indicates the absence of a spike and "n" the occurrence of n spikes) of window length Dt; thus, each word contained Dt/dt bins. We then tabulated the probability of occurrence $p_{Dt,dt}(W_i)$ of each word W_i and computed the associated entropies,

$$-\sum_{W_i} p_{Dt,dt}(W_i)\log_2 p_{Dt,dt}(W_i),$$

using W_i drawn from the repeated (Figure 4b) and random (Figure 4c) trials to estimate T and C, respectively (see also Strong et al., 1998).

Accurate estimation of the information using the direct method depends on the choice of Dt and dt and the firing rate. The low firing rates in our data (12 spikes/s on average) allowed us to obtain accurate estimates from as few as 50 trials. Most of our estimates were obtained using between 50 and 200 3 s trials for each (total and conditional entropy) condition. The information estimates were insensitive to the bin width (we used 4 and 8 ms bins) and usually maintained the same value up to a window size of ∼100 ms. In order to ensure that we were not overestimating information, we also studied the behavior of the information estimate as a function of the data set size. We found that the estimate remained robust until the data set for single cells was reduced to 50 trials, consistent with a theoretical worst-case (maximum entropy) estimate based on the number of trials needed for the homogeneous Poisson process.

Since the window sizes we used were not exact multiples of the refresh period (16.7 ms), the direct method captured information about the temporal phase of the refresh in each window. Similarly, the direct method could have also captured information concerning the spatial phase of the Gabor patch, which was independent of the stimulus motion. Information about phase is not "extensive"; i.e., it does not grow in proportion to the window size. In fact, for all windows longer than the refresh period, the available phase information per window is the same. In addition, we suspect the information rates we observed were actually underestimates of the maximum possible rate attainable, which could best be measured with a stimulus that changes more rapidly than the temporal precision of the neuron (∼3 ms).

Reconstruction Method
The reconstruction method involved a nonlinear algorithm to reconstruct the times when the stimulus moved in the cell's preferred direction from each spike train. At each frame, the stimulus either jumped in the preferred or antipreferred direction of the cell, so we represented both the original stimulus, s, and its reconstruction, $s_{est}(z)$, as a string of ones and zeros for each frame; ones indicated forward jumps, and zeros indicated backward jumps. For this method, we exclusively used the second mode of stimulation described above, where each 3 s stimulus presentation was different.

The algorithm worked as follows. First, we summed spikes occurring in a window centered some delay after the presentation of each stimulus frame. This delay could be thought of as the latency from the retina to the cell being studied and was close to 60 ms for all cells. Spikes were weighted by a Gaussian centered in the window, and the sum was thresholded to produce the reconstruction; when the sum was greater than the threshold, the reconstruction was set to one in that frame; if not, it was set to zero. For most cells, we used a stimulus distribution that did not allow two forward jumps in a row (i.e., the stimulus had a "refractory period"). In these cases, we sequentially searched for any consecutive pairs of ones in the reconstruction and set the reconstruction to zero in the frame with the smaller prethreshold sum. We then optimized the threshold, delay, and width of the Gaussian weighting function for each cell using gradient ascent in the mutual information using the formula below.

The mutual information between the stimulus s and the spike train z is bounded from below by the information between the stimulus and the reconstruction, which can be written as:

$$I(s,z) \geq I(s,s_{est}) = -\sum_{s_{est}} p(s_{est})\log_2 p(s_{est}) + \sum_s p(s) \sum_{s_{est}} p(s_{est}|s)\log_2 p(s_{est}|s).$$

The first term on the right hand side of this equation is the entropy of the estimate, which we approximated by assuming that forward jumps in the stimulus estimate were uncorrelated. This turned out to be a good approximation because (1) as the direct method above showed us, all cells produced spike trains that were consistent with a modulated Poisson process, and (2) the reconstruction had at most a third as many forward jumps as the original stimulus even for the most responsive cells, so any correlations in the stimulus due to its refractory period were on too short a timescale to affect the reconstructions. The second term on the right hand side of this equation is the entropy of the errors in our reconstruction, which we bounded from above by assuming that all errors were uncorrelated.

Signal Detection Theory versus Information Theory
The estimated information rates may be interpreted in terms of the familiar "discrimination power," which is given by the number of stimuli a system can discriminate with 100% reliability. In the two-stimulus case, the information available in a signal about these

stimuli maps directly onto the area of the ROC of signal detection theory (Green and Swets, 1966). The mapping is one to one for unimodal probability distributions. This can be readily seen by considering the fact that both the ROC area and relative entropy for unimodal distributions are strictly increasing functions of the distance between the modes. For example, in this case, 1 bit corresponds to perfect (100% correct) discrimination performance, whereas 0 bits correspond to chance performance (50% correct). Hence, information may reflect the reliability of discrimination performance. However, while information theory generalizes naturally to the analysis of signals with multimodal density functions, the analysis of such signals within the framework of conventional signal detection theory is awkward. Thus, information theory can be seen as a generalization of ROC analysis within which both reliable and unreliable discriminability can be readily assessed, regardless of the complexity of the probability distributions and the number of stimulus choices. For a complementary approach to discriminability, see Gabbiani et al. (1996).

LFP Analysis

In order to quantitate the relation between the spike count and the immediately preceding LFP, we defined a measure R as follows. The measure was defined only for the responses to the repeated stimulus, for which stimulus-locked increases in firing rate were well defined for many trials (Figure 4a). As illustrated in Figure 7a, associated with each reversal from the antipreferred to the preferred direction is an increase in the firing rate one latency (60 ms) later. Let T_i denote the time of the ith reversal of the stimulus motion (measured from the onset of the stimulus). Further, let N_{iq} represent the number of spikes on the qth trial in a 30 ms window following the latency (i.e., from $T_i + 60$ and $T_i + 90$), and let L_{iq} represent the average LFP on the qth trial in a 25 ms window before stimulus-triggered response (i.e., from $T_i + 25$ and $T_i + 50$). (Trials on which any "stray" spikes contaminated the prerespons period $[0, T_i]$ were excluded from this analysis; this was typically <10% of trials). N_{iq} was typically 0, 1, or 2 spikes. We then defined the corresponding average over trials of L_{iq} as

$$\bar{L}_i = \frac{1}{M}\sum_{q=1}^{M} L_{iq},$$

where M is the number of trials. We also defined $\bar{L}_i^0$ as the conditional average LFP for trials on which no spike occurred during the response period,

$$\bar{L}_i^0 = \frac{1}{M_0}\sum_{q0=1}^{M_0} L_{iq0},$$

where the index q_0 runs over only those trials on which $N_{iq} = 0$, and M_0 is the number of such trials. Finally, $Q_i = 1$ if $\bar{L}_i^0 > \bar{L}_i$; otherwise, $Q_i = 0$. Then, R is defined for each cell as

$$R = \frac{1}{K}\sum_{i=1}^{K} Q_i,$$

where K is the total number of reversals.

   The measure R so defined has the following properties. If the LFP associated with each stimulus reversal is uncorrelated with trial-to-trial variations in the spike count, then R = 0.5. If the average LFP preceding each stimulus-triggered event is always (i.e., for all events) higher on trials with no spikes than on trials with spikes, then R = 1. Note that R only assesses the trial-to-trial correlation between the spike count and the preceding LFP and not systematic variations in the LFP or the spike count due to the details of the stimulus.

References

Adrian, E.D. (1926). The Basis of Sensation: the Action of the Sense Organs (New York: W. W. Norton).

Albright, T.D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. J. Neurophysiol. *52*, 1106–1130.

Allen, C., and Stevens, C.F. (1994). An evaluation of causes for unreliability of synaptic transmission. Proc. Natl. Acad. Sci. USA *91*, 10380–10383.

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science *273*, 1868–1871.

Bair, W., and Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. Neural Comp. *8*, 44–66.

Bair, W., Cavanaugh, J.R., and Movshon, J.A. (1997). Reconstructing stimulus velocity from neuronal responses in area MT. In Advances in Information Processing Systems, Volume 9. Proceedings of Conference on Advances in Neural Information Processing Systems (Cambridge, MA: MIT Press), pp. 34–40.

Berry, M.J., Warland, D.K., and Meister, M. (1997). The structure and precision of retinal spike trains. Proc. Natl. Acad. Sci. USA *94*, 5411–5416.

Bryant, H.L., and Segundo, J.P. (1976). Spike initiation by transmembrane current: a white-noise analysis. J. Physiol. *260*, 279–314.

Cover, T.M., and Thomas, J.A. (1991). Elements of Information Theory (New York: John Wiley and Sons).

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

de Ruyter van Steveninck, R.R., Lewen, G.D., Strong, S.P., Koberle, R., and Bialek, W. (1997). Reproducibility and variability in neural spike trains. Science *21*, 1805–1808.

DeWeese, M. (1996). Optimization principles for the neural code. In Advances in Neural Information Processing, Volume 8. Proceedings of Conference on Advances in Neural Information Processing Systems (Cambridge, MA: MIT Press), pp. 281–287.

Dobkins, K.R., and Albright, T.D. (1994). What happens if it changes color when it moves? Neurophysiological experiments on the nature of chromatic input to macaque area MT. J. Neurosci. *14*, 4854–4870.

Dobrunz, L., and Stevens, C.F. (1997). Heterogeneity of release probability, facilitation, and depletion at central synapses. Neuron *18*, 995–1008.

Ferster, D., and Spruston, N. (1995). Cracking the neuronal code. Science *270*, 756–757.

FitzHugh, R. (1957). The statistical detection of threshold signals in the retina. J. Gen. Physiol. *40*, 925–948.

Gabbiani, F., Metzner, W., Wessel, R., and Koch, C. (1996). From stimulus encoding to feature extraction in weakly electric fish. Nature *384*, 564–567.

Gray, C. (1994). Synchronous oscillations in neuronal systems: mechanisms and functions. J. Comput. Neurosci. *1*, 11–38.

Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics. (New York: John Wiley and Sons).

Gur, M., Beylin, A., and Snodderly, D.M. (1997). Response variability of neurons in primary visual cortex (V1) of alert monkeys. J. Neurosci. *17*, 2914–2920.

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

Mainen, Z.F., and Sejnowski, T.J. (1995). Reliability of spike timing in neocortical neurons. Science *268*, 1503–1506.

Maunsell, J.H., and Van Essen, D.C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. J. Neurophysiol. *49*, 1127–1147.

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. Nature *341*, 52–54.

Optican, L.M., and Richmond, B.J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. J. Neurophysiol. *57*, 162–178.

Reich, D.S., Victor, J.D., Knight, B.W., Ozaki, T., and Kaplan, E. (1997). Response variability and timing precision of neuronal spike trains in vivo. J. Neurophysiol. *77*, 2836–2841.

Rieke, F., Warland, D., de Ruyter van Steveninck, R.R., and Bialek, W. (1997). Spikes. A Bradford Book (Cambridge, MA: MIT Press).

Shadlen, M.N., and Newsome, W.T. (1995). Is there a signal in the noise? Curr. Opin. Neurobiol. *5*, 248–250.

Shannon, C.E. (1948). A mathematical theory of communication. Bell Sys. Tech. J. *27*, 379–423.

Softky, W., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. J. Neurosci. *13*, 334–350.

Stevens, C.F., and Zador, A.M. (1996). Information through a spiking neuron. In Advances in Neural Information Processing Systems, Volume 8. Proceedings of Conference on Advances in Neural Information Processing Systems (Cambridge, MA: MIT Press), 75–81.

Stratford, K.J., Tarczy-Hornoch, K., Martin, K.A., Bannister, N.J., and Jack, J.J. (1996). Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. Nature *382*, 258–261.

Strong, S.P., Koberle, R.R., de Ruyter van Steveninck, R.R., and Bialek, W. (1998). Entropy and information in neural spike trains. Phys. Rev. Lett. *80*, 197.

Teich, M.C., Johnson, D.H., Kumar, A.R., and Turcott, R.G. (1990). Rate fluctuations and fractional power law noise recorded from cells in the lower auditory pathway of the cat. Hearing Res. *46*, 41–52.

Tolhurst, D.J., Movshon, J.A., and Dean, A.F. (1986). The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res. *23*, 775–785.

Tovee, M.J., Rolls, E.T., Treves, A., and Bellis, R.P. (1993). Information encoding and the responses of single neurons in the primate temporal visual cortex. J. Neurophysiol. *70*, 640–654.

Zador, A. (1998). The impact of synaptic unreliability on the information transmitted by spiking neurons. J. Neurophysiol. *19*, 1230–1238.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143.