Decoding Visual Information From a Population of Retinal Ganglion Cells

DAVID K. WARLAND, PAMELA REINAGEL, AND MARKUS MEISTER
Molecular and Cellular Biology Department, Harvard University, Cambridge, Massachusetts 02138

Warland, David K., Pamela Reinagel, and Markus Meister. Decoding visual information from a population of retinal ganglion cells. *J. Neurophysiol.* 78: 2336–2350, 1997. This work investigates how a time-dependent visual stimulus is encoded by the collective activity of many retinal ganglion cells. Multiple ganglion cell spike trains were recorded simultaneously from the isolated retina of the tiger salamander using a multielectrode array. The stimulus consisted of photopic, spatially uniform, temporally broadband flicker. From the recorded spike trains, an estimate was obtained of the stimulus intensity as a function of time. This was compared with the actual stimulus to assess the quality and quantity of visual information conveyed by the ganglion cell population. Two algorithms were used to decode the spike trains: an optimized linear filter in which each action potential made an additive contribution to the stimulus estimate and an artificial neural network trained by back-propagation to match spike trains with stimuli. The two methods performed indistinguishably, suggesting that most of the information about this stimulus can be extracted by linear operations on the spike trains. Individual ganglion cells conveyed information at a rate of 3.2 ± 1.7 bits/s (mean $\pm$ SD), with an average information content per spike of 1.6 bits. The maximal possible rate of information transmission compatible with the measured spiking statistics was 13.9 ± 6.3 bits/s. On average, ganglion cells used 22% of this capacity to encode visual information. When a decoder received two spike trains of the same response type, the reconstruction improved only marginally over that obtained from a single cell. However, a decoder using an ON and an OFF cell extracted as much information as the sum of that obtained from each cell alone. Thus cells of opposite response type encode different and nonoverlapping features of the stimulus. As more spike trains were provided to the decoder, the total information rate rapidly saturated, with 79% of the maximal value obtained from a local cluster of just four neurons of different functional types. The decoding filter applied to a given neuron's spikes within such a multiunit decoder differed substantially from the filter applied to that same neuron in a single-unit decoder. This shows that the optimal interpretation of a ganglion cell's action potential depends strongly on the simultaneous activity of other nearby cells. The quality of the stimulus reconstruction varied greatly with frequency: flicker components below 1 Hz and above 10 Hz were reconstructed poorly, and the performance was optimal near 2.5 Hz. Further analysis suggests that temporal encoding by ganglion cell spike trains is limited by slow phototransduction in the cone photoreceptors and a corrupting noise source proximal to the cones.

INTRODUCTION

A fundamental question in neuroscience is how the information relevant to behavior is represented in the activity of neurons (for review, see Abbott 1994; Bialek and Rieke 1992; Perkel and Bullock 1968; Rieke et al. 1997). The present work investigates the neural code employed by retinal ganglion cells in transmitting visual information to the brain. How do the spike trains of optic nerve fibers convey the visual scene projected on the retina? At this stage of the visual system, questions regarding the neural code can be phrased and answered particularly precisely for the following reasons: the ganglion cells are the only neurons transmitting visual information to the brain; the only variable they encode is the time-varying image on the retina; this stimulus can be controlled experimentally using well-developed technology for generating images; finally, the activity of multiple retinal ganglion cells can be monitored experimentally without damaging the retinal circuitry.

There have been many thorough investigations of how individual retinal ganglion cells respond to light (for review, see Dowling 1987; Shapley and Lennie 1985; Stone 1983). Typically one chooses a simple test stimulus, such as a flashing spot or a traveling grating, repeats the stimulus many times, and determines the time course of the ganglion cell's firing rate. Such studies have shown that one can clearly distinguish different functional types in the ganglion cell population and have characterized the sensitivity of the ganglion cell response depending on the temporal, spatial, and spectral composition of the light stimulus.

By comparison, we understand very little about how a population of these neurons collectively represents a complex visual scene and how subsequent stages of the visual system might extract features of the scene from this population activity. Therefore, the present work had two goals: to study the neural code at the level of a population of ganglion cells, rather than single neurons, and to analyze it from the point of view of the receiver of nerve messages. This amounts to answering the following question: given a certain pattern of spikes on optic nerve fibers, what is the likely visual scene that produced this activity?

In visual research, this approach to the neural code was pioneered by FitzHugh (1957, 1958). A cat retinal ganglion cell was driven with dim flashes of light near the response threshold. Then a ''statistical analyzer'' was applied to the ganglion cell spike train whose binary output reported whether a flash occurred or not. The performance of the analyzer was evaluated by the fraction of flashes detected correctly. This task was solved efficiently by a decoder that simply counted the number of spikes in a 30-ms window and reported a flash whenever this count exceeded a threshold.

This binary detection approach is appropriate for stimuli near the sensory threshold, where the retina can at best report the presence or absence of light. However, most of biological

vision happens far above threshold. In this regime, the retinal output can discriminate many different possible stimuli. Furthermore, the natural environment presents to the eye a very large ensemble of different inputs. This suggests an extension of FitzHugh's approach: constructing a decoder that estimates not only the presence or absence of light, but *how much* light there was at various times in the past. This amounts to reconstructing the visual stimulus sequence that led to the observed neural activity.

Such a reconstruction method was implemented by Bialek and colleagues (Bialek et al. 1991) to analyze how visual motion is encoded in the spike trains of neuron H1 in the fly's lobula plate. The fly's retina was presented with a large ensemble of stimuli from a randomly moving grating pattern. A decoder was constructed to estimate the velocity of the pattern from the recorded H1 spike train. In analogy to the binary detection task, the performance of the decoder was evaluated by analyzing its errors, namely the difference between the estimated and the actual stimulus. One such measure of performance is the Shannon information (Shannon and Weaver 1963) contained in the reconstruction about the true stimulus.

Although the fidelity of the neural code could be assessed in other ways, this information measure has enjoyed increasing popularity (de Ruyter van Steveninck and Laughlin 1996; Heller et al. 1995; Laughlin 1981; Rieke et al. 1995, 1997; Theunissen and Miller 1991). One reason is that, owing to its definition, information can only decrease in the course of signal processing. Therefore, if the output of a man-made decoder of spike trains conveys a certain amount of information about the stimulus, then the spike trains themselves must contain at least as much information. Furthermore, the structure of the decoder will reveal how to extract this knowledge. Another advantage is that the information measure abstracts from the specifics of the underlying communications process: information has dimensionless units. Thus one can compare the performance of the neural code across neurons, stimulus modalities, sensory systems, and species.

On this background, the present study was designed as follows (Fig. 1): To simplify subsequent analysis, we focused on the encoding of temporal aspects of the stimulus. The isolated retina of a tiger salamander was stimulated with a uniform gray field whose intensity varied randomly in time. Spike trains were recorded simultaneously from many ganglion cells using a multielectrode array. Then a multineuronal decoding algorithm was designed that took a collection of spike trains as input and produced an estimate of the time course of the light intensity.

Clearly it is essential that the decoding algorithm have the ability to extract the neural information, and a poor choice of decoder would compromise subsequent conclusions. Therefore, two very different decoding strategies were compared. One consisted of a multineuronal linear filter, a direct extension of the linear strategy used in previous reconstruction studies with single neurons (Bialek et al. 1991; Rieke et al. 1993; Roddey and Jacobs 1996). The other used an artificial neural network trained by back-propagation to match a given set of spike trains with the preceding stimulus. Remarkably, the two methods performed very similarly. The quality of the decoder's reconstruction was found to vary greatly depending on the number and types of neurons whose spike trains were used. Analysis of this relationship revealed to what extent ganglion cells carry redundant or independent information. Finally, we analyzed which stimulus features were best represented in the spike trains. In particular, it was found that the reconstruction quality depended strongly on flicker frequency, and we propose an explanation for the form of this relationship.

METHODS

Preparation and recording

Experiments followed the procedures described by Meister et al. (1994). Briefly, retinae from dark-adapted larval tiger salamanders (*Ambystoma tigrinum*) were isolated under infra-red illumination into oxygenated Ringer medium. A piece 3–4 mm in diameter was cut and placed ganglion cell side down in a superfusion chamber whose bottom contained a microelectrode array. The retina was superfused with oxygenated Ringer medium at 22°C. Action potentials were recorded from 61 electrodes distributed over an area of ~0.5 mm diameter. Typically, spikes from 20 to 30 individual ganglion cells could be distinguished and recorded continuously for an experiment lasting several hours.

Stimulation

The retina was stimulated by focusing an image of a computer monitor onto the photoreceptor layer. The screen was uniformly gray and modulated in intensity. The average intensity in each experiment was 7 mW/m^2 at the retina. The wavelength spectrum was as illustrated in Brainard (1989). For the salamander's red cone photoreceptor, the mean intensity produced an equivalent photon flux at the peak absorption wavelength (λ_{max} = 630 nm) of 9,500 photons/μm^2/s.

For diagnostic purposes, each experiment began with a brief period of stimulation using square-wave flashes with a period of 2 s. This was followed by 2–6 h of random flicker stimulation.
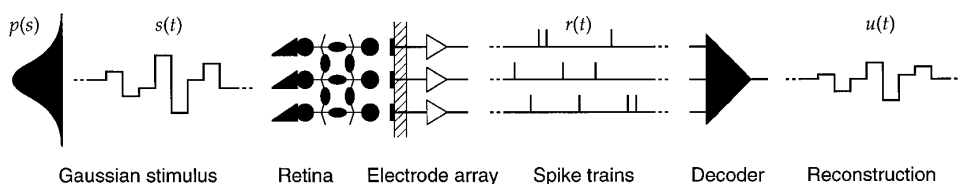

FIG. 1. Schematic overview of experiments and analysis. Gaussian random flicker of intensity $s(t)$ is projected onto the retina. Spike trains, $r^\nu(t)$, recorded from multiple ganglion cells, are presented to a decoding algorithm that extracts an estimate of the stimulus, $u(t)$.

For this purpose, the screen intensity was updated once every 15 ms or every 90 ms, by drawing a new value from a Gaussian probability distribution. The standard deviation of the Gaussian distribution, that is the root-mean-square contrast of the stimulus, was 35 or 24% of the mean. One experiment was performed using a binary stimulus distribution with the intensity drawn from only two possible values.

Ganglion cell types

For each ganglion cell, the reverse correlation of its spike train to the random flicker stimulus was computed (Meister et al. 1994). From the shape of this function—in particular its time-to-peak, τ—and the responses to flash stimulation, four types of light responses could be distinguished (Meister et al. 1995; M. Meister, unpublished data): fast OFF (typical $\tau \approx 57$ ms), slow OFF ($\tau \approx 72$ ms), weak OFF ($\tau \approx 82$ ms, with ON-OFF behavior under high-contrast light steps), and ON ($\tau \approx 107$ ms). Fast-OFF cells are further distinguished by their unusually large action potentials and form the most numerous class in our recordings. As observed previously, only ~20% of recorded neurons were ON cells. For all cell classes, the responses to this flicker stimulus appear to be rectified strongly (Berry et al. 1997). The same response properties have been observed on many other occasions (Smirnakis et al. 1997).

The detailed calculation of visual information rates reported here is based on 35 ganglion cells from three preparations. A fourth experiment, using the binary stimulus distribution, yielded different absolute information rates but very similar results regarding the high-frequency roll-off of information transmission and the interactions between different cells.

Optimal linear decoder

For numerical analysis, the stimulus intensity was scaled to zero mean and unit variance and discretized at time intervals of $\Delta t = 15$ ms.

$$s_i = \text{stimulus value in } i\text{th time interval } [i\Delta t, (i + 1)\Delta t]$$
$$= \frac{I_i - M}{C} \qquad (1)$$

where I_i is the intensity in the ith time interval, M is the mean intensity, and C is the standard deviation of the intensity. Similarly, the response was discretized by counting spikes in the corresponding time intervals

$$r_i^\nu = \text{number of spikes from cell } \nu \text{ in } i\text{th time interval} \qquad (2)$$

An estimate of the stimulus was obtained from the ganglion cell responses by convolving each spike train with a linear filter and adding the results along with a constant offset term

$$u_i = \text{stimulus estimate in } i\text{th time interval}$$
$$= a + \sum_\nu \sum_{j=0}^{N-1} r_{i-j}^\nu f_j^\nu \qquad (3)$$

where a is the constant offset, f_j^ν is the value of the decoding filter for cell ν at time $j\Delta t$ before the spike, and $N\Delta t$ is the length of the decoding filters. Each spike from cell ν makes an additive

contribution to this stimulus estimate reaching a period $N\Delta t$ back in time. The time course of this contribution is given by the values of the decoding filter, f_j^ν. The length of this filter, $N\Delta t$, was chosen empirically as 0.96 s: at times earlier than this the computed decoding filters were close to zero. To obtain higher frequency resolution in calculations of information spectra (see below), we sometimes used filters of length 3.84 s.

For compact notation, a vector formalism is useful for these relationships. Define a response matrix, $\mathbf{R}$, and vectors for the stimulus, $\mathbf{s}$, and reconstruction, $\mathbf{u}$

$$\mathbf{R} = \begin{bmatrix} 1 & r_0^1 & r_1^1 & \cdots & r_{N-1}^1 & \cdots & r_0^\nu & \cdots & r_{N-1}^\nu \\ 1 & r_1^1 & r_2^1 & \cdots & r_N^1 & \cdots & r_1^\nu & \cdots & r_N^\nu \\ 1 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & r_{M-1}^1 & r_M^1 & \cdots & r_{N+M-2}^1 & \cdots & r_{M-1}^\nu & \cdots & r_{N+M-2}^\nu \end{bmatrix}$$

$$\mathbf{s} = \begin{bmatrix} s_0 \\ s_1 \\ \cdots \\ s_{M-1} \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_0 \\ u_1 \\ \cdots \\ u_{M-1} \end{bmatrix} \qquad (4)$$

where

$$(N + M - 1)\Delta t = \text{duration of the recording} \qquad (5)$$

Then the reconstruction is obtained as

$$\mathbf{u} = \mathbf{R} \cdot \mathbf{f} \qquad (6)$$

where the transpose, $\mathbf{f}^{\mathrm{T}}$, of the multineuron decoding filter, $\mathbf{f}$, is given by

$$\mathbf{f}^{\mathrm{T}} = [a \quad f_0^1 \quad f_1^1 \cdots f_{N-1}^1 \cdots f_0^\nu \cdots f_{N-1}^\nu] \qquad (7)$$

This filter is optimized to minimize the squared difference between the stimulus and the reconstruction over the course of the experiment, $(\mathbf{s} - \mathbf{u})^{\mathrm{T}}(\mathbf{s} - \mathbf{u})$. The solution is obtained as

$$\mathbf{f} = (\mathbf{R}^{\mathrm{T}}\mathbf{R})^{-1} \cdot (\mathbf{R}^{\mathrm{T}}\mathbf{s}) \qquad (8)$$

Note this is an analytic result with no free parameters. The term $\mathbf{R}^{\mathrm{T}}\mathbf{s}$ corresponds to the reverse correlation between the stimulus and the spike trains. The term $\mathbf{R}^{\mathrm{T}}\mathbf{R}$ contains correlation functions among the spike trains. Thus the statistics of the spike trains and the stimulus completely determine the optimal decoding filter.

The limiting case of very sparse spike trains is instructive: if the spikes are spaced more than $N\Delta t$ apart, then the correlation matrix $\mathbf{R}^{\mathrm{T}}\mathbf{R}$ is diagonal and each cell's decoding filter is equal to the cell's spike-triggered average stimulus. This can be understood because, in this limit, each spike provides a message about the stimulus independent of that of all other spikes. The estimate of the preceding stimulus that minimizes the squared error is simply the mean of the stimulus distribution conditional on a spike. In general, however, many spikes are observed within one integration time, particularly if several neurons are included in the analysis. In this case, $\mathbf{R}^{\mathrm{T}}\mathbf{R}$ alters the shape of the reverse correlation to produce the optimal decoding filter.

Finally, given stimulus $\mathbf{s}$ and response $\mathbf{R}$, the optimal reconstruction is found to be

$$\mathbf{u} = \mathbf{R} \cdot \mathbf{f} = \mathbf{R}(\mathbf{R}^{\mathrm{T}}\mathbf{R})^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{s} \qquad (9)$$

Power spectra

The power spectrum of the stimulus was computed as follows. The stimulus, s_i, was divided into nonoverlapping blocks of length $N\Delta t$, equal to the length of the decoding filters. Each block was Fourier transformed, without data windowing, to yield

$$\tilde{s}_j = \text{Fourier transform of stimulus at frequency } j/2N\Delta t \qquad (10)$$

At each frequency, the squared modulus of the Fourier coefficient was averaged over all blocks. Power at negative frequencies was added to the value at the corresponding positive frequency, producing the ''one-sided'' power spectrum

$$P_j^{(S)} = \text{power spectral density of stimulus at frequency } j/2N\Delta t$$
$$= \langle |\tilde{s}_j|^2 + |\tilde{s}_{-j}|^2 \rangle_{\text{blocks}} \qquad (11)$$

Power spectra of the reconstruction, $P_j^{(U)}$, and of the reconstruction error, $P_j^{(E)}$, were computed in the same fashion, by Fourier transforming u_i and $(u_i - s_i)$, respectively.

Transmitted information

To assess the overall quality of the reconstruction, we determined how much information it contained about the visual stimulus. Let **s** and **u** denote short segments of length $N\Delta t$ of the stimulus and the reconstruction, respectively. Following Shannon (Rieke et al. 1997; Shannon and Weaver 1963), the mutual information between **s** and **u** corresponds to the reduction in uncertainty about **s** obtained by observing **u**

$$I_{\mathbf{s},\mathbf{u}} = H_{\mathbf{s}} - H_{\mathbf{s}|\mathbf{u}} \qquad (12)$$

where $H_{\mathbf{s}}$ = entropy of the stimulus = $-\sum_{\mathbf{s}} p(\mathbf{s}) \log_2 p(\mathbf{s})$, $p(\mathbf{s})$ = a priori probability of **s** in the stimulus ensemble, $H_{\mathbf{s}|\mathbf{u}}$ = conditional entropy of the stimulus given knowledge of **u** = $-\sum_{\mathbf{u}} p(\mathbf{u}) \sum_{\mathbf{s}} p(\mathbf{s}|\mathbf{u}) \log_2 p(\mathbf{s}|\mathbf{u})$, $p(\mathbf{u})$ = probability of **u**, and $p(\mathbf{s}|\mathbf{u})$ = probability of **s** given knowledge of **u**. In the present experiments, **s** was drawn from a Gaussian distribution, independently in subsequent time bins, and thus

$$p(\mathbf{s}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{S}}} e^{-(1/2)\mathbf{s}^T\mathbf{S}^{-1}\mathbf{s}} \qquad (13)$$

where **S** = covariance matrix of the stimulus = $\langle \mathbf{ss}^T \rangle$. Whereas the a priori probability distribution of the stimulus, $p(\mathbf{s})$, is known by construction, the conditional distribution, $p(\mathbf{s}|\mathbf{u})$, must be measured. This is difficult to achieve in full generality, but a partial measure is given by the second moment of the reconstruction error

$$\mathbf{E} = \text{covariance matrix of the reconstruction error}$$
$$= \langle (\mathbf{s} - \mathbf{u})(\mathbf{s} - \mathbf{u})^T \rangle \qquad (14)$$

Given the second moment, the probability distribution with the maximal entropy is a Gaussian (Cover and Thomas 1991). Thus one obtains an upper bound on the conditional entropy of the stimulus, $H_{\mathbf{s}|\mathbf{u}}$, by approximating the conditional distribution with a Gaussian

$$p(\mathbf{s}|\mathbf{u}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{E}}} e^{-(1/2)(\mathbf{s}-\mathbf{u})^T\mathbf{E}^{-1}(\mathbf{s}-\mathbf{u})} \qquad (15)$$

This leads to a lower bound on the information conveyed by the stimulus reconstruction

$$I_{\mathbf{s},\mathbf{u}} = H_{\mathbf{s}} - H_{\mathbf{s}|\mathbf{u}} > \log_2 \sqrt{\det (\mathbf{SE}^{-1})} \qquad (16)$$

This expression is valid in any orthonormal basis of stimulus space, but it is evaluated most easily in the frequency domain. In this basis, **s** and **u** contain the Fourier components of the stimulus and reconstruction. The covariance matrices **S** and **E** are diagonal due to time translation invariance, and their elements correspond to the power spectra of the stimulus and the reconstruction error, respectively. Thus one finds

$$I_{\mathbf{s},\mathbf{u}} > \sum_{j=0}^{N/2} I_j = \sum_{j=0}^{N/2} \log_2 (P_j^{(S)}/P_j^{(E)}) \qquad (17)$$

where I_j = information spectral density at frequency $j/2N\Delta t$ = $\log_2 (P_j^{(S)}/P_j^{(E)})$.

In practice, the above sum over frequencies was truncated at 20 Hz because higher frequencies made little systematic contribution. Furthermore, this estimate of mutual information was corrected for the finite duration of the experiment. For this purpose, a second decoding filter was computed with the same number of elements, N, but constrained to positive times after the action potential. Because this decoder effectively was asked to predict the future of the stimulus, one expects no systematic relationship between stimulus and prediction. Correspondingly, the mutual information was very small, typically 0.002 bits/s. This confirms that the recordings were sufficiently long to average over random coincidences between stimulus and spike trains and that the decoding filters truly reflect neural function. The information in the prediction was subtracted from the information in the reconstruction to obtain a lower bound on the information extracted from the spike trains.

Capacity

To assess the efficiency of information transmission, we estimated the capacity of individual neuronal spike trains. This amounts to the maximal information transmission rate such a neuron could sustain and is limited by its spiking statistics, specifically the entropy rate of the spike train (MacKay and McCulloch 1952; Shannon and Weaver 1963). To estimate the entropy, successive interspike intervals were taken to be independently generated symbols, which leads to

$$H_{\Delta t} = \text{entropy per unit time of a spike train binned at a resolution of } \Delta t$$

$$= -\frac{\sum_n p_n \log_2 p_n}{\sum_n p_n n \Delta t} \qquad (18)$$

where p_n = probability of an interspike interval of $n\Delta t$. This estimate ignores any possible correlations among nearby interspike intervals, and thus represents an upper bound on the spike train entropy (Rieke et al. 1993). Because the above methods produced

a lower bound on transmitted information (*Eq. 17*), the ratio of information rate to capacity

$$\epsilon = \text{efficiency of transmission}$$

$$= \frac{\text{information rate}}{\text{capacity}} = \frac{I}{H} \qquad (19)$$

is estimated by a lower bound.

Both the spike train entropy and the information rate depend on the time resolution Δt with which the spike trains are binned. As Δt increases, the entropy rate decreases, because many discriminable spike intervals are lumped together in the same value. Similarly, the information rate decreases once Δt exceeds the timing accuracy of the action potentials (Rieke et al. 1997). Under the present stimulus conditions, salamander ganglion cells fire spikes whose time of arrival jitters with a standard deviation of 5–10 ms across identical stimulus repeats (Berry et al. 1997). Correspondingly, we measured the entropy and the information rate for a range of bin sizes between 4 and 16 ms. Over this range, the coding efficiency varied very little, by only 11% on average over 14 cells, suggesting that this is indeed close to the intrinsic timing accuracy of the spike trains. All subsequent calculations of information rate and capacity were performed with $\Delta t = 15$ ms.

Artificial neural networks

Artificial neural networks were constructed and trained using the University of Toronto Simulator (UTS) libraries (Department of Computer Science, University of Toronto, Toronto, ON M5S 1A4, Canada; freely accessible at ftp.cs.toronto.edu:pub/xerion/). Three-layer, fully connected, feed-forward networks with ten hidden units were used, as illustrated in Fig. 2.

The input layer to the network was presented with a window of spike trains 0.96 s long, binned at 15 ms resolution. Each input unit represented the number of spikes in one time bin of one cell's spike train. The value of the output unit represented the stimulus estimate for the first time bin of the input window. In addition to the weighted outputs of the preceding layer, each hidden unit and the output unit also received a bias weight. For a given sample of the spike trains, the activity of the various units was given by

$$r_i^\nu = \text{number of spikes from ganglion cell } \nu \text{ in time bin } i$$

$$a_j = \text{activity of hidden unit } j = f\left(\sum_{\nu,i} r_i^\nu w_{ij}^\nu + w_{0j}\right)$$

$$w_{ij}^\nu = \text{hidden unit weights}$$

$$w_{0j} = \text{bias weight}$$

$$f(x) = \frac{1}{1 + e^{-x}}$$

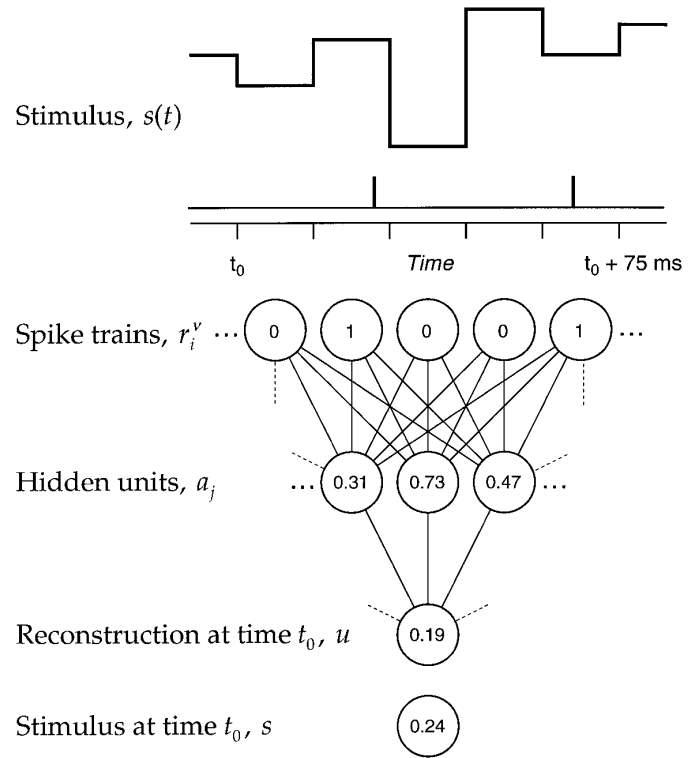$$u = \text{activity of output unit} = f\left(\sum_j a_j v_j + v_0\right)$$

FIG. 2. Artificial neural network used to decode the ganglion cell spike trains. See description in text.

$$v_j = \text{output unit weights}$$

$$v_0 = \text{bias weight} \qquad (20)$$

Weights were initialized to random values chosen from a Gaussian distribution with a mean of 0 and a standard deviation inversely proportional to the number of weights in the layer. Three different random initial conditions were used for each network. The networks were trained by back-propagation (Rumelhart et al. 1986) to minimize the mean squared error (MSE) between the stimulus estimate and the true stimulus. The direction of the change to each weight was computed using conjugate gradient descent (cgRudi in UTS), and the size of the step was computed by line search, which provides an adaptive learning rate (lsRay in UTS). The network's performance was tested during training by reconstructing the stimulus for part of the recording not used to train the network. Training was terminated when the MSE for this test segment stopped decreasing. Finally, the network's performance was assessed using a third data set reserved for validation. The reported measurements of mutual information between the stimulus and its reconstruction are from this validation data set.

Several nets were developed with variations on these methods, which affected the time required for training but produced no difference in decoding performance. Specifically, fully connected nets were trained with fewer or more hidden units (0–25); sparsely connected nets were trained with five hidden units dedicated to each ganglion cell and five hidden units dedicated to each coarse time window; the activation function in *Eq. 20* was changed to $f(x) = \tanh(x)$; the error measure was changed to the adaptive sum square (MSE normalized to the variance of the noise in each run); and nets were compared with and without
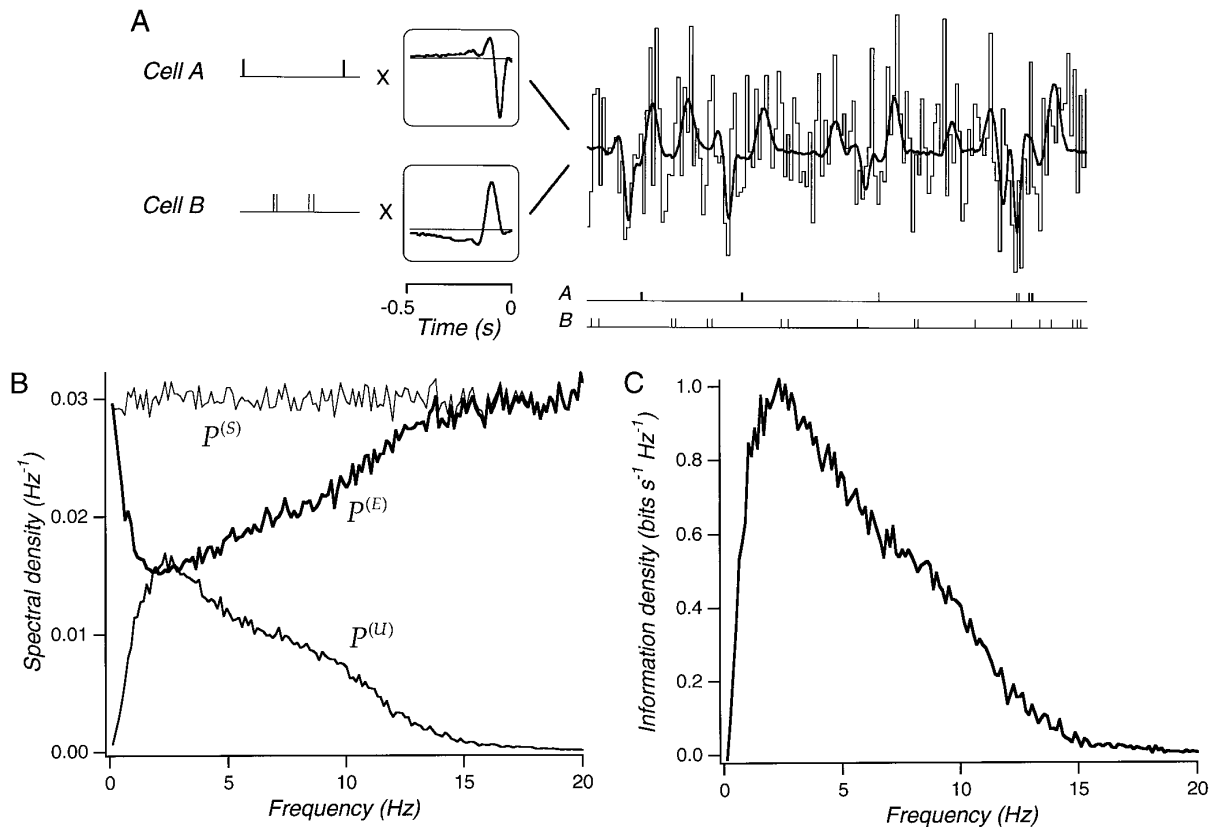
FIG. 3.   Linear decoding of 2 ganglion cell spike trains. *A*: 2 ganglion cell spike trains (brief segments in *left* and *bottom* of *right panel*) are convolved with their respective decoding filters (*middle*) and summed to yield the stimulus estimate (*right*, thick trace) to be compared with the real stimulus (*right*, thin trace). Decoding filters were 3.84 s long, and only the 1st 0.5 s is shown. *B*: power spectral density as a function of frequency for the real stimulus, $P^{(S)}$, the reconstruction, $P^{(U)}$, and the reconstruction error, $P^{(E)}$. *C*: information spectral density of this reconstruction as a function of frequency.

weight decay (scale 0.001 in UTS) and with and without injected noise.

RESULTS

A linear multineuronal decoder

To assess how a local population of ganglion cells represented a random flicker stimulus, we attempted to reconstruct the intensity time course from the recorded spike trains alone. The first method employed a linear filter, extending the single-neuron methods pioneered by Bialek et al. (1991) to an array of cells. During stimulus reconstruction, each spike from a ganglion cell adds a contribution to the stimulus estimate that extends a short period into the past. The shapes of the various cells' filter functions are adjusted to minimize the mean-squared difference between the stimulus estimate and the true stimulus over the duration of a long experiment (*Eq. 8*). The decoding process is illustrated in Fig. 3*A,* which shows two typical ganglion cell spike trains, the temporal profiles of the optimal decoding filters for these cells, and the stimulus reconstruction obtained by passing the spike trains through the two-cell filter.

DECODING FILTERS.   The shapes of the decoding filters provide a measure of each cell's visual message. For a few tens of milliseconds before the action potential (which occurs at *time 0*), the decoding filter makes no contribution to the stimulus estimate because of the finite response latency of the ganglion cell. Due to delays in phototransduction and subsequent neural processing, the occurrence of a spike conveys no information about the immediately preceding stimulus. Similarly, the filter function vanishes at very early times far preceding the spike. The intervening period of ~0.5 s reflects the ganglion cell's integration time. Within this interval, both cells' filters have a biphasic waveform. This suggests that their firing was dependent on changes in the light intensity rather than the absolute level. The two waveforms differ in sign: one of these cells was triggered preferentially by an increase, the other by a decrease in light intensity.

SPECTRAL ANALYSIS OF THE RECONSTRUCTION.   The reconstruction in Fig. 3*A* appears to capture large and slow variations in the stimulus reasonably well, but fails to track it during periods of rapid flicker. To assess more quantitatively what fraction of the stimulus was reconstructed at each flicker frequency, we calculated the power spectral density of the stimulus, the reconstruction, and the reconstruction

error (see METHODS). Figure 3*B* shows that the power in the stimulus was fairly constant over the range 0–20 Hz. However, the power in the reconstruction dropped off sharply at high frequencies. For this cell pair, the 50% roll-off point is at 9 Hz. At frequencies >15 Hz, the reconstruction power essentially vanished and thus the decoder extracted no information about the stimulus.

Interestingly, the reconstruction power also dropped off sharply at low frequencies. This was not caused by the limited length of the decoding filters: the filter function is essentially zero at times before −0.5 s and tests with longer filter functions did not improve the reconstructions. At low frequencies, the 50% roll-off was near 1 Hz. The peak power in the reconstruction was found at 2.5 Hz.

ESTIMATE OF INFORMATION TRANSMISSION. By comparing the stimulus reconstruction to the true stimulus, one can obtain an estimate of how much information the decoder extracted from the spike trains (Bialek et al. 1991, 1993; Rieke et al. 1997). As derived in METHODS, a lower bound on this information rate is given by the power spectra of the stimulus and the reconstruction error

$$I = \text{extracted information rate}$$
$$= \int_0^\infty I(f)\,df \tag{21}$$

where $I(f)$ = information spectral density at frequency $f >$ $\log_2 [P^{(S)}(f)/P^{(E)}(f)]$, $P^{(S)}(f)$ = power spectral density of the stimulus, and $P^{(E)}(f)$ = power spectral density of the reconstruction error. Figure 3*C* shows the information density $I(f)$ derived from the same two-cell reconstruction. As observed above for the reconstruction power, the information rate dropped sharply at frequencies <1 Hz and >9 Hz with a peak information rate at 2.5 Hz. Overall, the decoder extracted from these two cells ≥7.5 bits/s of information about the visual stimulus. Such information density curves will be used in subsequent sections to summarize the performance of various decoders.

Redundancy

To explore whether ganglion cells carry independent or redundant information about the stimulus, we compared the reconstruction obtained from a pair of cells with the reconstructions based on either cell alone. Because the stimulus was spatially uniform, each ganglion cell received the same input. Thus if a decoder using two cells recovers more information than decoders using either cell alone, this reflects the extent to which the cells differ in their encoding of the same stimulus.

Figure 4*A* shows the information density curves for two fast-OFF cells. Note that the reconstruction using both spike trains simultaneously performed only slightly better than that using either spike train alone. The information recovered by the two-cell decoder was much less than the sum of the information recovered by the two single-cell decoders. One
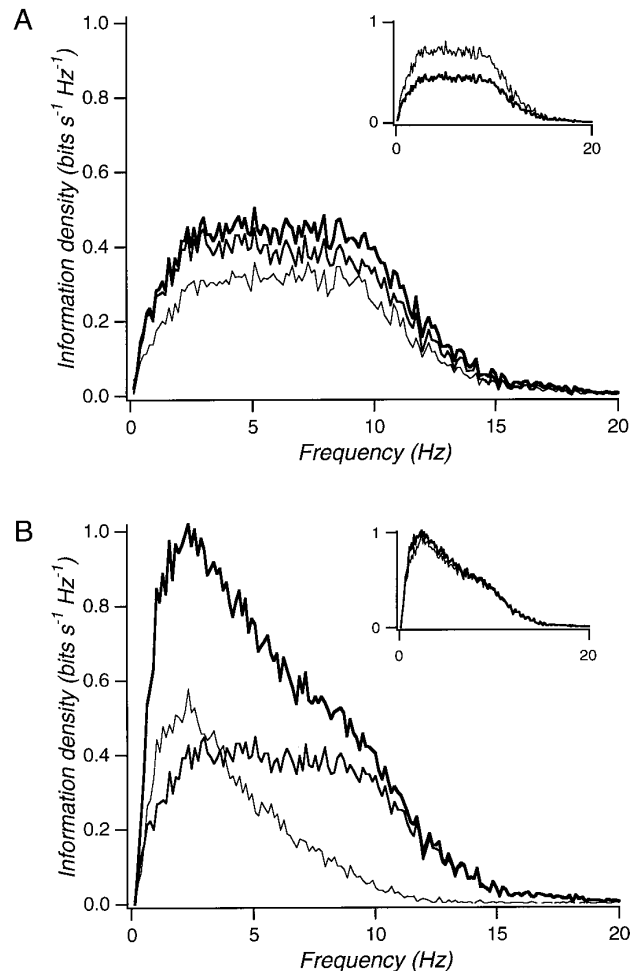

FIG. 4. Redundancy between 2 spike trains. *A*: 2 OFF-type cells, named *A* and *H*. Information density vs. frequency, derived from decoding only cell *H* (I_H, thin trace), only cell *A* (I_A, medium trace), and both cells (I_{AH}, thick trace). *Inset* compares the sum of the information in the 2 single-cell reconstructions ($I_A + I_H$, thin trace) to the information obtained from the 2-cell reconstruction (I_{AH}, thick trace). *B*: an OFF cell (*cell A*) and an ON cell (*cell B*). Information density vs. frequency plotted as in *A*. I_B, thin trace; I_A, medium trace; I_{AB}, thick trace. *Inset* compares $I_A + I_B$ (thin trace) to I_{AB} (thick trace).

concludes that these two cells transmitted redundant information.

Figure 4*B* shows the information density curves derived from a fast-OFF cell and an ON cell. Note that the fast-OFF cell conveys information to significantly higher frequencies, consistent with the longer integration time of the ON response (see METHODS and the filters in Fig. 3*A*). When the two responses were combined, they provided significantly more information about the stimulus than either cell taken separately. In fact, the information density recovered from the two spike trains was essentially equal to the sum of the information densities recovered from each cell. This means that ON cells and OFF cells differ not only by the sign of their response to light, but rather each encodes aspects of the stimulus that are not covered at all by the other cell.

Saturation of temporal information

As Figs. 3 and 4 show, the two-cell decoders reconstruct only a fraction of the true stimulus. There is ample room for improvement, particularly at frequencies >10 Hz and <1 Hz. How would the quality of the stimulus estimate improve as the spike trains from more and more cells are included?

Figure 5A shows a brief segment of stimulus and spike trains for 14 ganglion cells (named A–N). By successively adding more spike trains (in alphabetic order of cell name), 14 different linear decoders and their stimulus reconstructions were computed. As expected, the spectrum of the information density increased monotonically as more and more cells were included in the reconstruction (Fig. 5B). However, this curve changed very little for reconstructions using more than four cells.

The total information rate summed over all frequencies reached a plateau of ~10 bits/s as more cells were included for the reconstruction (Fig. 5C). The exact shape of this curve depended somewhat on the order in which cells were added, though the final plateau was independent of this order. Because neurons of the same functional type tended to carry redundant information, we began by combining cells with recognizably different light responses (see METHODS). As seen in Fig. 5C, combining one cell each from the four different response types already provided >79% of the maximal information. When the decoder was restricted to cells of the fast-OFF functional type, 10 neurons provided only 20% more information than a single cell. Apparently, a handful of ganglion cells can convey all the information available about this stimulus.

Spike meaning depends on the context from other cells

As derived in METHODS, the optimal decoding filter for a given neuron depends not only on the response properties of that cell, but also on the responses of other cells included in the same decoder. Figure 6A shows that these effects can be very strong. It compares the filter function for an OFF-type cell in several different decoders: from one using only that cell's spike train, to one using eight additional spike trains as well. The shape of the optimal filter function changed substantially in the context of responses from these other neurons. In particular, the amplitude decreased and the time course became far more oscillatory. Neurons with very similar single-cell decoders often showed strikingly different filters within a multicell decoder (Fig. 6B). This implies that the optimal interpretation of an action potential from one ganglion cell depends strongly on the messages received from other neurons. Thus temporal information appears to be distributed across several neurons and only can be recovered fully by processing their signals simultaneously.

Coding efficiency

Clearly, the reconstruction of the stimulus obtained from the best multineuron linear decoder does not recover all of the visual stimulus, and it misses large amounts of information at very low and very high frequencies. Does this reflect a deficiency of the decoder or is the information simply not contained in the spike trains? It appears useful to determine whether the missing information could, in principle, be carried by the ganglion cell signals. Do the measured firing rates and firing statistics allow the transmission of more information than has been extracted by the decoder?

The capacity of a spike train, that is the highest rate at which it could transmit information, can be estimated by monitoring its variability during a long period of time while it is encoding many different stimuli. A very regular spike train, for example one with a constant interspike interval, can encode little or no information because variations in the stimulus cannot translate into variations in the spike train. The formal information-theoretic measure of such variability is the entropy of the spike train. In the present analysis we estimated an upper bound on this entropy, H, from the variation in the interspike intervals (see METHODS). Once the entropy of the spike train, H, is known, together with the rate at which it carries visual information, I, one can evaluate how efficiently the spike train is used by

$$\epsilon = \text{coding efficiency} = \frac{I}{H} \qquad (22)$$

Figure 7 illustrates results from this analysis for a population of fast-OFF ganglion cells. Among different cells, the information rate varied over more than an order of magnitude, with an average of 3.7 ± 1.5 bits/s (mean $\pm$ SD). Similarly, the spike train entropy varied over a great range with an average of 14.4 ± 5.9 bits/s. Remarkably, both the visual information and the entropy rates varied almost proportionally to the mean firing rate, with an average entropy of 6.6 bits/spike and information of 1.9 bits/spike. Thus the coding efficiency was almost constant across the population of fast-OFF cells with values clustered near 26%. The same relationship was found in three different preparations and under various stimulus conditions. On average over all cells analyzed, the efficiency was 22%. Thus these ganglion cells used ~1/4 of their spike train capacity to encode the visual stimulus. Although this conclusion is subject to a possible overestimation of the entropy (see METHODS), it appears likely that the statistical structure of a ganglion cell spike train does not significantly restrict its visual information content.

Artificial neural networks

Finally, one is led to consider whether the information missing from the stimulus reconstructions is in fact represented in the spike trains but cannot be accessed by a linear decoder. It might be encoded in a form such that an optimal estimate of the stimulus requires a very nonlinear combination of different action potentials, for example, the recognition of a particular pattern of many spikes.

In principle, this problem could be approached systemati-

cally by expanding the stimulus estimate into higher and higher powers of the response spike trains (Bialek et al. 1991). In such a Taylor series, *Eq. 6* represents only the
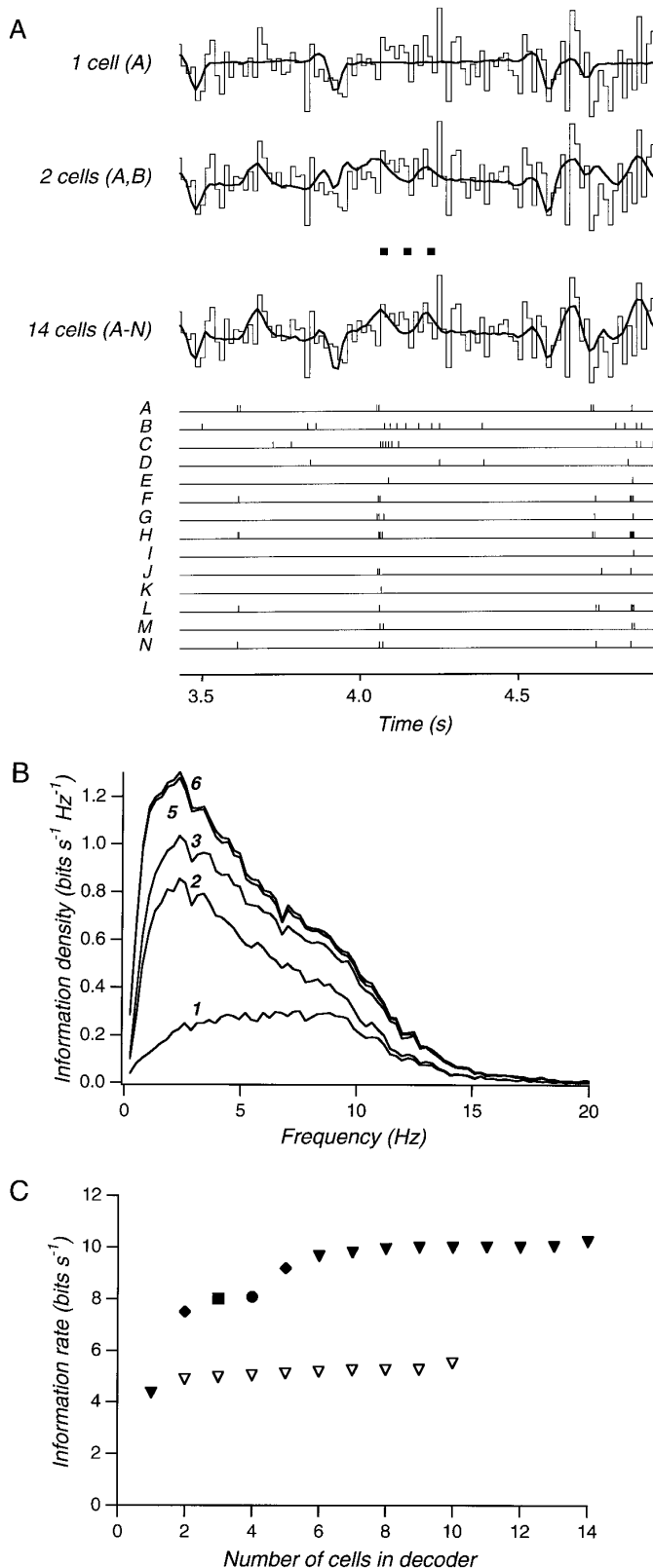
terms of 0th and 1st order; all subsequent terms would contribute a nonlinear reconstruction. However, the number of coefficients in these higher order filters explodes rapidly. When many spike trains are to be combined, this problem is further amplified, and it is not practical to calculate even the most general second-order filter from the available data. As an alternative approach to finding a nonlinear decoder, we have trained artificial neural networks (hereafter neural networks) to perform the decoding task. This allowed exploration of a wide space of possible nonlinear decoders, although the algorithm for exploring that space is not guaranteed to find the best solution.

A segment of all ganglion cell spike trains was presented to the input layer of a feed-forward neural network with three layers (Fig. 2, *Eq. 20*). The network output consisted of a single unit whose value was taken as the stimulus estimate at the beginning of the spike train window. Both at the hidden layer and in the convergence to the final output unit, signals were combined nonlinearly through a sigmoidal activation function. The weights in the network were trained by back-propagation to minimize the mean squared error between the estimated and the true stimulus. Qualitatively this procedure entailed presenting a spike train segment to the network, determining the error in the output, determining the amount by which each weight should be changed to decrease the error, and repeating this for all time segments on multiple cycles through the data set until the network was trained fully (see METHODS for details).

Figure 8*A* compares the stimulus reconstructions obtained from the fully trained neural network and the optimal linear filter, using spike trains from 14 cells in the experiment of Fig. 5*A*. Note that the two reconstructions are virtually superimposable. No systematic deviations could be recognized between the two stimulus estimates.

To further assess the performance of the ANN, we computed the mutual information between its stimulus reconstruction and the real stimulus using the same approach applied to the linear reconstruction (see METHODS and Fig. 5). As the number of spike trains used by the decoder was increased from 1 to 14, the neural network recovered almost the same amount of information as the optimal linear filter (Fig. 8*B*). Under all conditions tested, the information rates derived from the neural network and the linear filter differed by, at most, 10%. We conclude that the neural network and the linear decoder are indistinguishable in performance, with regard to both the quantity and the qualitative features of the visual information extracted from ganglion cell spike trains.

One might expect such a tight correspondence if the neural

FIG. 5. Visual information from a population of ganglion cells. *A*: brief segment of the stimulus (*top*, thin trace), the spike trains of 14 retinal ganglion cells (*bottom*), and the stimulus reconstructions from various multicell decoders (*top*, thick trace). *B*: information density derived by adding successive ganglion cells to the decoder, using *cell A* alone (*trace 1*), *A* and *B* (*2*), *cells A–C* (*3*), up to *cells A–F* (*6*). *C*: total information derived as successive ganglion cells are added to the decoder. Symbol shape denotes the response type of the last ganglion cell added: ▼, fast OFF; ◆, ON; ■, weak OFF; and ●, slow OFF. Open symbols show the information density derived from decoding only spike trains from cells of the fast OFF type.
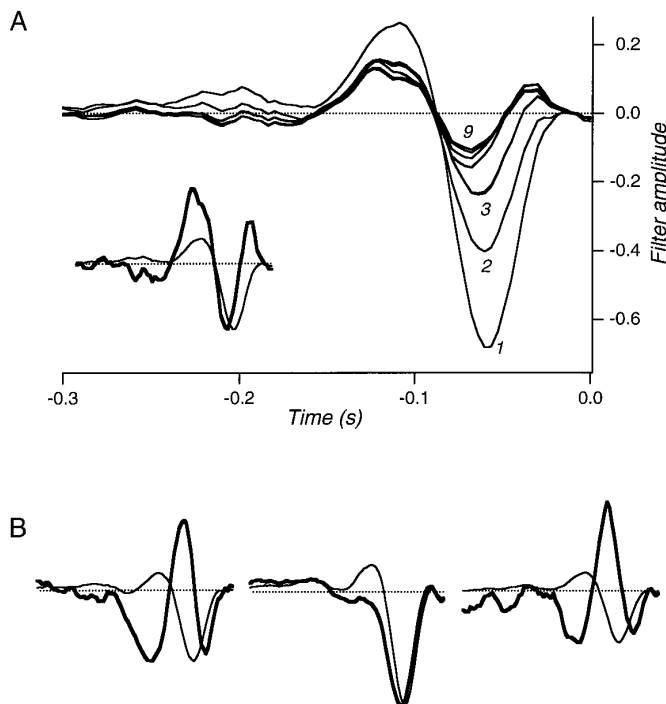
FIG. 6. Dependence of spike meaning on context from other cells. *A*: filter function for an OFF cell (*cell A* in Fig. 5*A*) in a single-cell decoder (*trace 1*) and in decoders using additional cells (*2–9*). *Inset*: traces *1* (thin line) and *9* (thick line) scaled to the same peak. *B*: same comparison for 3 other OFF cells, displayed as in the *inset* of *A*.

network effectively acted as a linear filter, for example, by only using the linear portion of the activation curve (*Eq. 20*). Instead, it was found that the hidden units of the neural network explored their entire nonlinear response range over the ensemble of input data. Figure 8*C* shows the distribution of output values of the 10 hidden units for a neural network processing 14 spike trains whose output is shown in Fig. 8*A*. Comparison with the activation function shows that all the hidden units produced outputs in a nonlinear part of the range for some of the spike train examples. Thus the neural network transformed the spike train nonlinearly to a compressed representation in the hidden layer, yet the final computation was equivalent to a linear transformation. Clearly, the design of the neural nets did not constrain them to a linear regime.

Finally, we tested whether these neural networks could in practice extract nonlinear spike codes that were inaccessible to a linear decoder. For this purpose, artificial spike trains were generated with known coding properties. Three spike trains were simulated, two ON-type and one OFF-type, whose instantaneous firing rate varied as a linear function of the stimulus (Fig. 9). No attempt was made to emulate the detailed spiking statistics of ganglion cells, but the mean firing rates and visual integration times were chosen to match typical neurons. The response function was chosen so that the firing rate did not saturate; therefore, all the stimulus information available in the spike trains was encoded linearly. These three spike trains were presented for a stimulus reconstruction to a linear decoder and to a neural network,

and the stimulus information contained in the reconstructions was computed as above. As expected, the linear filter and neural network extracted very similar amounts of information from these simulated spike trains (Fig. 9).

To produce a nonlinear code for the same stimulus, two spike trains were constructed by adding the spikes of the OFF cell to each of the ON cell spike trains. This mimics two ON-OFF ganglion cells, whose OFF signal derives from a shared input neuron. The optimal interpretation of these two spike trains requires the decoder to recognize that synchronous spikes from the two cells mean something very different from solitary spikes in only one cell; in fact, these two events signal stimulus episodes of the opposite sign. These distinct interpretations cannot be achieved by a linear filter decoder. Operating on these two spike trains, the neural network extracted about twice as much information as the linear filter (Fig. 9). This proves that the neural network decoder can significantly outperform the linear decoder when the stimulus is represented in a very nonlinear fashion.

In these simulations, the neural network did not recover all the information originally present in the three separate spike trains, in part because occasional spurious coincidences of spikes from the two ON cells create an unresolvable ambiguity. Moreover, the neural network failed to outperform a linear decoder on some other contrived nonlinear codes that we tested in these simulations. Thus the training algorithm is not guaranteed to find the optimal interpretation of spike trains, but the network decoder can clearly adapt to a broader variety of codes than the linear filter.

When operating on a population of real ganglion cell spike trains, however, the performance of neural network and linear filter was always indistinguishable even though the neural network did explore its nonlinear operating range. Thus it appears that the linear decoder has access to most of the stimulus information contained in ganglion cell spike trains and that the deficiencies of the stimulus reconstruction at low and high flicker frequencies are not due to the simple architecture of the linear decoder.
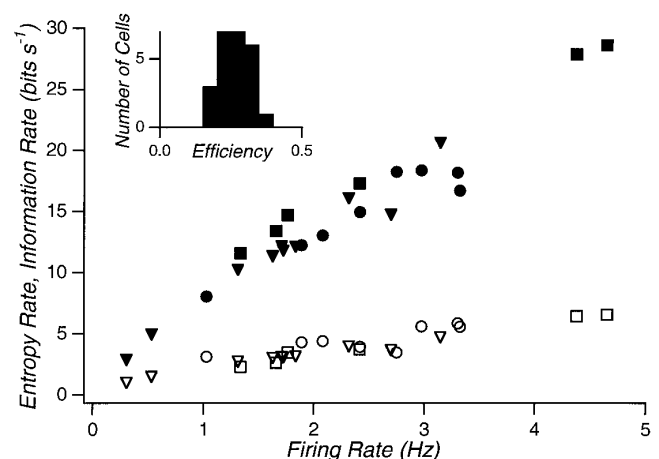

FIG. 7. Coding efficiency. Entropy rate (closed symbols) and decoded information rate (open symbols) for individual fast-OFF cells plotted as a function of the mean firing rate. Stimulus parameters in the 3 experiments were: 15-ms update interval and 35% contrast (triangles and squares); 90-ms update interval and 24% contrast (circles). *Inset*: histogram of the coding efficiency, namely the ratio of information rate to entropy.
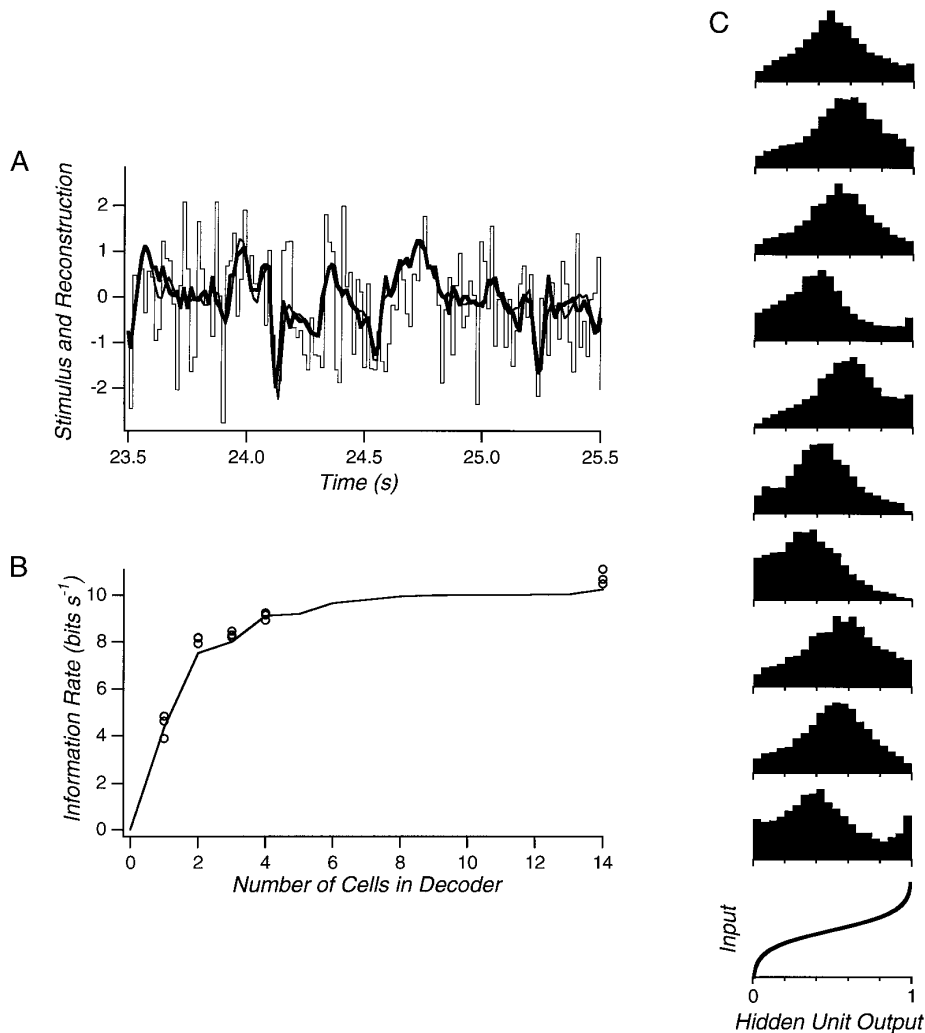
FIG. 8. Comparison of 2 types of decoder. *A*: brief segment of the reconstructions of the stimulus (thin line) from the spike trains of 14 cells, obtained by the optimal linear decoder (medium line), and a fully trained artificial neural network (thick line). *B*: total decoded information obtained by the neural network (○) and the optimal linear decoder (——) as successive ganglion cells were added to the decoder. For each set of cells, 3 networks were trained from different initial weight settings. *C*: histograms of activity levels for the 10 hidden units in the neural network that produced the reconstruction in *A*. *Bottom*: activation curve for each hidden unit to illustrate the limited range of linear operation.

DISCUSSION

Stimuli and their reconstruction

Our experiments employed a rather simple visual stimulus: a uniform gray field whose intensity varied randomly in time. This focus on temporal processing served to limit the complexity of the analysis. In the first study of neural coding by simultaneous spike trains, it seemed essential to keep the computations efficient so that we could survey a range of phenomena in multineuronal decoders. Furthermore under these conditions, all ganglion cells experienced the same stimulus, and it became feasible to compare directly the coding properties of different neurons, even if recorded in different parts of the retina. Nevertheless, many ganglion cells are driven more strongly by a flickering checkerboard than by a uniform field of the same temporal contrast owing to their antagonistic receptive field profile (Smirnakis et al. 1997). Thus one certainly expects further insights from an analysis of responses to spatially varying stimuli, and the associated technical difficulties are being tackled.

We assessed how well the visual stimulus was represented in the spike trains by explicitly reconstructing the time course of the intensity from the spike trains. This does not imply an assumption that the salamander ultimately attempts to reconstruct the raw visual image. Instead, it provides a way to explicitly reveal which aspects of the stimulus are well encoded and which are not. Note, however, that the spike trains could encode stimulus features that are not at all useful for a reconstruction. For example, a cell might fire depending on the absolute value of intensity fluctuations. In absence of another cell that encodes the sign of the intensity fluctuation, that spike train can make no contribution to the stimulus reconstruction even though it obviously transmits visual information. Thus our estimates of visual information in the spike trains are by necessity lower bounds.

Information rates and coding efficiency

The absolute information rates achieved by single salamander ganglion cells are relatively low by comparison to other neural systems: ca 4 bits/s compared with 23 bits/s in frog auditory afferents stimulated with broadband sounds (Rieke et al. 1995), 64 bits/s in the fly's H1 neuron (Bialek
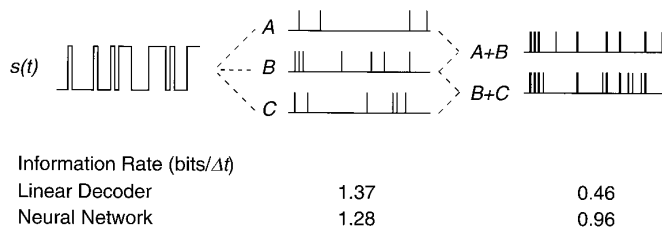
FIG. 9. Performance of the linear decoder and the artificial network on simulated ganglion cell spike trains. Stimulus consisted of binary flicker—as used during one of the experiments—with values $s_i = 0$ or 1, updated randomly every time interval Δt. Three spike trains (A–C) were simulated as follows: in the ith time bin, $[i\Delta t, (i + 1)\Delta t]$, a spike was generated with probability

$$p_i = 0.2 \begin{cases} \dfrac{1}{r - l} \displaystyle\sum_{j=i+l}^{i+r-1} s_j, & \text{for ON cells} \\[2ex] 1 - \dfrac{1}{r - l} \displaystyle\sum_{j=i+l}^{i+r-1} s_j, & \text{for OFF cells} \end{cases}$$

These simulated cells simply integrate the stimulus over the interval $[l\Delta t, r\Delta t]$ and modulate their firing probability linearly with the result. Parameters were chosen as follows: A, ON, $l = -8$, $r = -4$; B, OFF, $l = -4$, $r = -2$; C, ON, $l = -5$, $r = -2$. To produce a nonlinear code for the same stimulus, the spikes in the OFF-type spike train (B) were added into each ON-type spike train ($A + B$, $B + C$); spikes from B are marked with thick lines. Information per time interval Δt that the optimal linear filter and a fully trained artificial neural network extracted from each of these representations is given below.

et al. 1991), 155 bits/s in vibratory receptors of the bullfrog sacculus (Rieke et al. 1993), and 294 bits/s in cricket mechanoreceptors (Rieke et al. 1993). As argued below, retinal information transfer likely is limited by the slow process of phototransduction in the receptors. On the other hand, the average information content of a ganglion cell spike, ca 1.8 bits, is comparable with that found in other neural systems even though the total information rates vary dramatically: 0.66 bits for frog auditory afferents, 0.75 bits in the fly study (Rieke et al. 1997), 2.6 bits for bullfrog vibratory receptors, and 3.2 bits for the cricket mechanoreceptors. Thus an information content per spike of ~ 1–3 bits emerges as a constant across many different sensory systems. Similarly, the efficiency of information transmission, namely the fraction of a spike train's capacity used for coding, is comparable across systems: 22% averaged over all cells in the present study compared with 11% for frog auditory afferents, 30% in the fly (F. Rieke and D. K. Warland, unpublished data), and 50–60% in the bullfrog and cricket studies. Considering that these efficiency values are likely underestimates (see METHODS), it appears that the statistics of the spike trains, particularly their firing rates, generally are well matched to the needs for information transmission. However, it should be noted that the efficiency of the neural code can depend strongly on the stimulus ensemble used to evaluate it. Rieke et al. (1995) observed a fourfold to fivefold increase in both the information rate and the coding efficiency of frog auditory afferents when using sounds whose spectra were shaped like natural frog calls rather than white noise. It remains to be seen how retinal ganglion cells behave under stimuli with more naturalistic spatio-temporal statistics.

Interactions between neurons

Two cells of the same response type typically encoded redundant information, that is, the stimulus reconstruction did not improve much by monitoring two or more cells. This was a somewhat unexpected result, because it often is assumed that the brain must average over many noisy neural signals to obtain a reliable message. Instead, it appears that individual ganglion cells are sufficiently reliable, at least under the stimulus conditions employed here (Berry et al. 1997). Furthermore, as argued below, it appears that nearby ganglion cells share a limiting noise source, such that observing several spike trains of the same type does not improve knowledge of the stimulus. By comparison, ON and OFF cells generated mostly independent information. Thus a good stimulus reconstruction required the inclusion of only one ganglion cell from each recognizably different functional type. Nevertheless, it is known that some of these functional types have a high degree of receptive field overlap (Meister, unpublished data). A full understanding of the coding properties in such an arrangement will require an analysis with spatially varying stimuli (Warland and Meister 1995).

Within a decoder that monitored many spike trains simultaneously, the decoding filter associated with a given neuron was quite different from the corresponding filter in a single-neuron decoder. Thus the meaning of a cell's action potentials depends critically on the spikes from other nearby cells. Note that none of these interactions within the ganglion cell population could have been detected by recording from one cell at a time. Thus attempts to derive a ''population code'' from serial single-neuron recordings may well be incomplete (Georgopoulos 1990).

Limits to retinal information transfer

As more and more ganglion cell spike trains were recruited to reconstruct the visual stimulus, the information clearly saturated (Fig. 5C) even though the reconstruction was still lacking much of the true stimulus, particularly at high temporal frequencies exceeding 10 Hz (Fig. 5B). This deficiency was not due to a limitation of the ganglion cells' information capacity (Fig. 7). Furthermore, both linear and nonlinear decoders experienced this saturation (Fig. 8). Thus the information that cannot be reconstructed must have been discarded in the course of processing by the retinal circuit. What determines this strict limit to visual information? The steep roll-off at high flicker frequencies points to the slowest neuron in the circuit, namely the photoreceptor.

CONE FLASH RESPONSES. Figure 10A shows flash responses of a salamander cone at mean light intensities comparable with our measurements. The outer segment membrane current, measured in an isolated photoreceptor (Matthews et al. 1990), shows a monophasic time course. However, the membrane voltage, recorded in the intact retina (Pasino and Marchiafava 1976), has a much faster, biphasic flash response. This difference, which is more pronounced under uniform than local illumination, arises in part from the de-
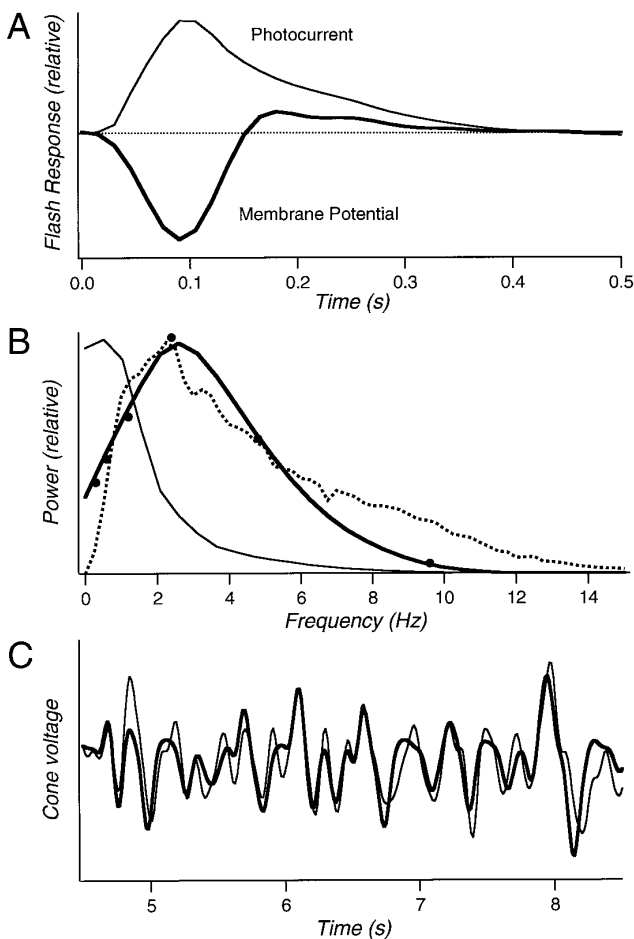
FIG. 10. Light response of salamander red cone photoreceptors, light-adapted to the mean intensity of our experiments. *A*: flash response of the outer segment current (thin trace) (from Matthews et al. 1990) and the membrane potential (thick trace) (transformed from the Bode plot in Pasino and Marchiafava 1976). *B*: power spectrum of the flash response of the outer segment current (thin trace), obtained from Fourier transform of the trace in *A*, and the membrane potential (thick trace), an interpolation through data of Pasino and Marchiafava (1976), shown with a filled circle. Equivalent signal-to-noise ratio of the best stimulus reconstruction (*Eq. 23*) is shown with dotted line. *C*: estimated membrane potential of a red cone, $v(t)$ (thin trace), and its reconstruction, $v_{est}(t)$ (thick trace), by the optimal 14-cell linear decoder of Fig. 5*A*.

layed inhibitory feedback that the cone receives from horizontal cells (Baylor et al. 1971).

If the cone transduces the flickering light stimulus linearly—a reasonable approximation given the moderate intensity contrast of 0.35—the output signal will again be Gaussian with a power spectrum given by the Fourier transform of the flash response. These spectra are plotted in Fig. 10*B*: clearly the high-frequency components of the stimulus are attenuated greatly, more so in the membrane current than in the membrane voltage.

RELATION TO THE RECONSTRUCTION QUALITY. The transmitted information depends not only on attenuation of the signal but also on the magnitude of the noise. If the output of the cone photoreceptor is corrupted by the addition of a Gaussian noise, then the summed signal transmits information about the stimulus at a spectral density of

$$I(f) = \log_2\left(1 + \frac{S(f)}{N(f)}\right) \tag{23}$$

where $S(f)$ is the power spectrum of the stimulus transduced by the receptor, $N(f)$ is the power spectrum of the noise, and all spectra are "one-sided," evaluated only at positive frequencies (Shannon and Weaver 1963). Using this relation and the information density of Fig. 5*B* (6 cells), we plot the signal-to-noise ratio (SNR), $S(f)/N(f)$, equivalent to the best stimulus reconstruction (Fig. 10*B*). One finds that this spectrum is very similar to the spectrum of the cone's membrane potential, $S(f)$; in particular the two have approximately the same high-frequency roll-off. This implies that the noise, $N(f)$, added downstream from the cones has a rather flat power spectrum.

Note that the spectrum of the cone's membrane current lies at much lower frequencies than that of the membrane voltage or the SNR. If the dominant noise source were within the cone, adding to the membrane current derived from phototransduction, its peak would have to be at even lower frequencies to produce the measured SNR spectrum. This would conflict with our current understanding of photoreceptor noise, which includes a component with the same spectrum as the flash response ("discrete noise") and components at higher frequencies ("continuous noise") (Baylor et al. 1980).

Even if the SNR for individual cones is low at high frequencies, one might still expect that the high-frequency signal could be extracted from the network noise: individual ganglion cells pool signals from many cones, and the multicell decoder can access the signals of many ganglion cells. The fact that this was not possible could be explained if the limiting noise source is shared among many ganglion cells such that pooling their signals no longer improves the SNR. This also is suggested by recent measurements of correlations between ganglion cell spike trains: nearby cells have a strong tendency to fire synchronously both with and without visual stimulation (Meister et al. 1995). These strong correlations may result from shared amacrine cell input. They are limited to an intercellular distance of about one ganglion cell receptive field diameter, and thus further visual information might be obtained by combining the spike trains from more distant cells. However, under natural conditions this is not an option because distant points on the retina receive different stimulation.

RECONSTRUCTION OF THE CONE SIGNAL. Given that the photoreceptors play such an important part in shaping the visual signal, one might ask how faithfully the ganglion cell spike trains represent the output of the cones. Assuming that the cone responds linearly under our stimulus conditions, its membrane potential $v(t)$ is obtained by convolving the stimulus time course, $s(t)$, with the flash response, $i(t)$, (Fig. 10*C*)

$$v(t) = \int s(t')i(t - t')dt' \tag{24}$$

What is the optimal linear reconstruction of this signal given the spike trains? Note that the stimulus reconstruction by

the optimal decoder $u(t)$ is a linear function of the true stimulus $s(t)$ (*Eq. 9*). Because $v(t)$ is a linear transform of $s(t)$, $v_{est}(t)$ is the same linear transform of $u(t)$. Thus the optimal reconstruction of the membrane potential $v_{est}(t)$ is given by convolving $u(t)$ with the cone flash response

$$v_{est}(t) = \int u(t')i(t - t')dt' \qquad (25)$$

Figure 10*C* shows a segment of the estimated membrane potential $v(t)$ and its reconstruction $v_{est}(t)$. It appears that the population of ganglion cell spike trains can encode the cone signal with remarkable fidelity even though the light response of an individual ganglion cell bears little resemblance to the cone response. Furthermore, this signal can be extracted from the spike trains by simple linear decoding.

Thus the observed high-frequency limit to retinal information transfer probably is imposed by the low-pass filtering of the phototransduction process in combination with a broadband noise source in subsequent retinal circuitry, which is shared by many retinal ganglion cells. At the low-frequency end, the roll-off of the information extracted from the spike trains is considerably steeper than in the cone response (Fig. 10*B*). This probably results from temporal differentiation of the signal by amacrine circuits in the inner retina (Maguire et al. 1989), such that the ganglion cell spike trains ultimately carry little or no information about the intensity of a constant light ($f = 0$ Hz).

Linear decoding of retinal signals

All of the information useful for a reconstruction of the stimulus appeared to be accessible with a linear decoder even though we worked hard to improve performance with a nonlinear algorithm. Others have reached similar conclusions; for example, in decoding the spike train of fly H1 neurons using a systematic expansion of the stimulus in powers of the response, the second-order term contributed $<5\%$ of the information in the linear term (Bialek et al. 1991). It has been suggested that this property of the neural code allows a simple processing strategy in subsequent circuits (Bialek et al. 1991; Rieke et al. 1997). Even though these neurons may not need to reconstruct the original stimulus, they would have access to particular features of that stimulus by simply filtering their input spike trains linearly. This can be achieved by convolving each afferent neuron's signal with the associated postsynaptic potential and summing over the dendritic tree.

In summary, even though the process by which the retina encodes the stimulus with spikes is very nonlinear, the inverse operation, by which the stimulus is extracted from the spikes, may be essentially linear. This is not a contradiction: consider, for example, a stimulus variable, s, being encoded by two response channels, r_+ and r_-, each a half-wave rectifier, $r_\pm = (|s| \pm s)/2$. Clearly the encoding process is nonlinear, whereas the reconstruction is obtained simply as $s = r_+ - r_-$. This example may partly account for the linearity of retinal decoding, if one identifies the two rectifiers with the ON and OFF channels of retinal processing. Again, it will be instructive to

see whether this property is borne out in the processing of spatially varying stimuli. There are indications that specific patterns of spikes across ganglion cells play a role in encoding spatial information (Meister 1996; Warland and Meister 1995), and their optimal interpretation may well require more than a linear decoder.

REFERENCES

ABBOTT, L. F. Decoding neuronal firing and modelling neural networks. *Q. Rev. Biophys.* 27: 291–331, 1994.
BAYLOR, D. A., FUORTES, M. G., AND O'BRYAN, P. M. Receptive fields of cones in the retina of the turtle. *J. Physiol. (Lond.)* 214: 265–294, 1971.
BAYLOR, D. A., MATTHEWS, G., AND YAU, K. W. Two components of electrical dark noise in toad retinal rod outer segments. *J. Physiol. (Lond.)* 309: 591–621, 1980.
BERRY, M. J., WARLAND, D. K., AND MEISTER, M. The structure and precision of retinal spike trains. *Proc. Natl. Acad. Sci. USA* 94: 5411–5416, 1997.
BIALEK, W., DEWEESE, M., RIEKE, F., AND WARLAND, D. Bits and brains: information flow in the nervous system. *Physica A* 200: 581–593, 1993.
BIALEK, W. AND RIEKE, F. Reliability and information transmission in spiking neurons. *Trends Neurosci.* 15: 428–434, 1992.
BIALEK, W., RIEKE, F., DE RUYTER VAN STEVENINCK, R. R., AND WARLAND, D. Reading a neural code. *Science* 252: 1854–1857, 1991.
BRAINARD, D. Calibration of a computer controlled color monitor. *Color Res. Appl.* 14: 23–34, 1989.
COVER, T. M. AND THOMAS, J. A. *Elements of Information Theory.* New York: Wiley, 1991.
DE RUYTER VAN STEVENINCK, R. R. AND LAUGHLIN, S. B. The rate of information transfer at graded-potential synapses. *Nature* 379: 642–645, 1996.
DOWLING, J. E. *The Retina: An Approachable Part of the Brain.* Cambridge, MA: Harvard Univ. Press, 1987.
FITZHUGH, R. The statistical detection of threshold signals in the retina. *J. Gen. Physiol.* 40: 925–948, 1957.
FITZHUGH, R. A statistical analyzer for optic nerve messages. *J. Gen. Physiol.* 41: 675–692, 1958.
GEORGOPOULOS, A. P. Neural coding of the direction of reaching and a comparison with saccadic eye movements. *Cold Spring Harb. Symp. Quant. Biol.* 55: 849–859, 1990.
HELLER, J., HERTZ, J. A., KJAER, T. W., AND RICHMOND, B. J. Information flow and temporal coding in primate pattern vision. *J. Comput. Neurosci.* 2: 175–193, 1995.
LAUGHLIN, S. A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch. [C]* 36: 910–912, 1981.
MACKAY, D. AND MCCULLOCH, W. S. The limiting information capacity of a neuronal link. *Bull. Math. Biophys.* 14: 127–135, 1952.
MAGUIRE, G., LUKASIEWICZ, P., AND WERBLIN, F. Amacrine cell interactions underlying the response to change in the tiger salamander retina. *J. Neurosci.* 9: 726–735, 1989.
MATTHEWS, H. R., FAIN, G. L., MURPHY, R. L., AND LAMB, T. D. Light adaptation in cone photoreceptors of the salamander: a role for cytoplasmic calcium. *J. Physiol. (Lond.)* 420: 447–469, 1990.
MEISTER, M. Multineuronal codes in retinal signaling. *Proc. Natl. Acad. Sci. USA* 93: 609–614, 1996.

MEISTER, M., LAGNADO, L., AND BAYLOR, D. A. Concerted signaling by retinal ganglion cells. *Science* 270: 1207–1210, 1995.

MEISTER, M., PINE, J., AND BAYLOR, D. A. Multi-neuronal signals from the retina: acquisition and analysis. *J. Neurosci. Methods* 51: 95–106, 1994.

PASINO, E. AND MARCHIAFAVA, P. L. Transfer properties of rod and cone cells in the retina of the tiger salamander. *Vision Res.* 16: 381–386, 1976.

PERKEL, D. H. AND BULLOCK, T. H. Neural coding. *Neurosci. Res. Program Bull.* 6: 221–348, 1968.

RIEKE, F., BODNAR, D. A., AND BIALEK, W. Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc. R. Soc. Lond. B Biol. Sci.* 262: 259–265, 1995.

RIEKE, F., WARLAND, D., AND BIALEK, W. Coding efficiency and information rates in sensory neurons. *Europhys. Lett.* 22: 151–156, 1993.

RIEKE, F., WARLAND, D., DE RUYTER VAN STEVENINCK, R. R., AND BIALEK, W. *Spikes: Exploring the Neural Code.* Cambridge, MA: MIT Press, 1997.

RODDEY, J. C. AND JACOBS, G. A. Information theoretic analysis of dynamical encoding by filiform mechanoreceptors in the cricket cercal system. *J. Neurophysiol.* 75: 1365–1376, 1996.

RUMELHART, D. E., HINTON, G. E., AND WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature* 323: 533–536, 1986.

SHANNON, C. E. AND WEAVER, W. *The Mathematical Theory of Communication.* Chicago: University of Illinois Press, 1963.

SHAPLEY, R. AND LENNIE, P. Spatial frequency analysis in the visual system. *Annu. Rev. Neurosci.* 8: 547–583, 1985.

SMIRNAKIS, S. M., BERRY, M. J., WARLAND, D. K., BIALEK, W., AND MEISTER, M. Adaptation of retinal processing to image contrast and spatial scale. *Nature* 386: 69–73, 1997.

STONE, J. *Parallel Processing in the Visual System.* New York: Plenum, 1983.

THEUNISSEN, F. E. AND MILLER, J. P. Representation of sensory information in the cricket cercal sensory system. II. Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. *J. Neurophysiol.* 66: 1690–1703, 1991.

WARLAND, D. K. AND MEISTER, M. Multi-neuronal firing patterns among retinal ganglion cells encode spatial information (Abstract). *Invest. Ophthalmol. Vis. Sci.* 36, *Suppl.*: 932, 1995.