JOURNAL OF NEUROPHYSIOLOGY
Vol. 72, No. 6, December 1994. *Printed in U.S.A.*

On Temporal Codes and the Spatiotemporal Response of Neurons in the Lateral Geniculate Nucleus

D. GOLOMB, D. KLEINFELD, R. C. REID, R. M. SHAPLEY, AND B. I. SHRAIMAN
AT&T Bell Laboratories, Murray Hill, New Jersey 07974; Mathematics Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892; Laboratory of Neurobiology, The Rockefeller University, New York, 10021; and Center for Neural Science, New York University, New York, New York 10003

SUMMARY AND CONCLUSIONS

1. The present work relates recent experimental studies of the temporal coding of visual stimuli (McClurkin, Optican, Richmond, and Gawne, *Science* 253: 675, 1991) to the measurements of the spatiotemporal receptive fields of neurons within the lateral geniculate of primate.

2. We analyze both new and previously described magnocellular and parvocellular single units. The spatiotemporal impulse response function of the unit, defined as the time-resolved average firing rate in response to a weak stimulus flashed at a given location and time, is characterized by the singular value decomposition. This analysis allows one to represent the impulse response by a small number, two to three, of spatial and temporal modes. Both magnocellular and parvocellular units are weakly nonseparable, with major and minor modes that account, respectively, for ~78 and 22% of the response. The major temporal mode for both types is essentially identical for the first 100 ms. At later times the response of magnocellular units changes sign and decays slowly, whereas the response of parvocellular units decays relatively rapidly.

3. The spatiotemporal impulse response function completely determines the response of a unit to an arbitrary stimulus when linear response theory is valid. Using the measured impulse response, combined with a rectifying neuronal input-output relation, we calculate the responses to a complete set of spatial luminance patterns constructed of "Walsh" functions. Our predicted temporal responses are in qualitative agreement with those reported for parvocellular units (McClurkin, Optican, Richmond, and Gawne, *J. Neurophysiol.* 66: 794, 1991). Under the additional assumptions of Poisson statistics for the probability of spiking and a plausible background firing rate, we predict the performance of a unit in the Walsh pattern discrimination task as quantified by mutual information. Our prediction is again consistent with the reported results.

4. Last, we consider the issue of temporal coding within linear response. For stimuli presented for fixed time intervals, the singular value decomposition provides a natural relation between the temporal modes of the neuronal response and the spatial pattern of the stimulus. Although it is tempting to interpret each temporal mode as an independent channel that encodes orthogonal features of the stimulus, successively higher order modes are increasingly unreliable and do not significantly increase the discrimination capabilities of the unit.

INTRODUCTION

The neurophysiological correlate of sensation is a change in the spike output rate of one or more neurons in response to a change in the pattern of external stimulation. A priori, the relation between the output of a neuron and features of external stimuli may be complex. In practice, this relation is often simple for neurons involved in early stages of sensory pathways. Important and well-studied examples occur in the mammalian visual system, in which neurons at early stages respond to input localized to a restricted region of space. This region is referred to as the receptive field (RF).

The RF of a neuron is a qualitative descriptor that is usually specified independently of features, such as luminance, orientation, size, and velocity, that affect the firing rate of the neuron. However, the description of the RF is clearly intertwined with that of feature selectivity, e.g., the shape of the RF will determine the orientation preference of the unit (e.g., Wörgötter and Koch 1991). Thus, in principle, a more general description of a unit can be formulated that allows one to predict the response of the unit to specific spatiotemporal input patterns. In practice, this description has been achieved only for neurons whose response is linear or dominated by a specific nonlinearity (see articles in Pinter and Nabet 1992).

The response of a neuron is said to be linear if it satisfies the principle of superposition, i.e., the combined response to different stimuli is equal to the sum of the responses to individual stimuli. Previous experiments on the visual system of cat and monkey suggest that the response of many neurons to weak stimuli through the level of the lateral geniculate nucleus (LGN) (Enroth-Cugell and Robson 1966; Hochstein and Shapley 1976; So and Shapley 1979) and possibly primary visual cortex (Jagadeesh et al. 1993; Jones and Palmer 1987; Movshon et al. 1978; Reid et al. 1987, 1991; Shapley et al. 1991) is linear to good approximation. Within the linear approximation, the structure of the receptive field of the cell can be fully described by the measured response to any complete set of stimuli. A particularly simple complete set is localized flashes, and the description that results from correlating the output of a neuron with the past location and time of a flash, i.e., so called reverse correlation (de Boer and Koyper 1968; Podvigin et al. 1974), is denoted as the spatiotemporal impulse response (STIR) function for the unit. Numerous investigators have used reverse correlation techniques to construct the STIR function of units in the LGN (Podvigin et al. 1974; Reid and Shapley 1992) and primary visual cortex (McLean and Palmer 1989; Palmer et al. 1991; Reid et al. 1987, 1991). Although knowledge of the STIR is sufficient to predict the

response of the unit to an arbitrary input provided that its output remains within the linear regime, the consequences of linearity on issues of coding features of the stimuli by the neuronal spike train have not been properly examined (but see Atick 1992; Bialek 1991).

An alternative description of feature selectivity is considered by Optican and Richmond and colleagues (McClurkin et al. 1991c, 1994). These authors present measurements of the temporal response of neurons at subcortical and cortical levels in the primate visual system (Gawne et al. 1991; McClurkin et al. 1991a–c, 1994; Optican and Richmond 1987; Richmond and Optican 1987, 1990; Richmond et al. 1987, 1990). They conclude that spatial aspects of a stimulus are coded in terms of the temporal structure of the neuronal response.

Motivated by the evidence that the neuronal response in early visual areas is close to linear, we reexamine the results of Optican and Richmond and colleagues on the temporal coding properties of units in the LGN (McClurkin et al. 1991a–c) in light of previous (Reid and Shapley 1992) and new measurements of the spatiotemporal structure of the receptive field for these units. We use the measured STIR function and the assumed rectifying nonlinearity of the neuronal input-output relation to compute the expected temporal response of our LGN units to the set of Walsh pattern stimuli used by McClurkin et al. (1991b). Following the latter authors, we compute the principal components of the temporal response. We find that the principal components predicted on the basis of measured STIR functions are in qualitative agreement with those observed by McClurkin et al. (1991a,b). We proceed with the comparison by computing the mutual information between the set of stimuli and the corresponding responses. Again, reasonable agreement with the results of McClurkin et al. (1991a) is found.

We conclude that the measurements of McClurkin et al. (1991b) are consistent with the linear response data of Reid and Shapley (1992 and this work). The temporal structure of the neuronal response to Walsh patterns, observed by the former investigators, originates in the temporal properties of the neuronal response to a brief local stimulus. As expected from the general principles of information theory, the characterization of the response that retains more of its temporal structure, e.g., a time-resolved rather than time-averaged characterization, carries greater mutual information. However, we express reservation with respect to the interpretation of the temporal principal components as "codes" of the spatial structure of the stimulus. The notion of a code appears redundant in the linear regime, where a well defined linear input-output relation exists. Furthermore, as our analysis will clarify (DISCUSSION), such a "code" would apply only to spatial stimuli with an identical, specific time course. An appealing alternative is to think of distinct modes of the response as independent information channels.

The outline of this paper is as follows. In METHODS we discuss the procedures used to acquire data. In RESULTS the measured STIR functions for magnocellular and parvocellular units of macaque LGN are presented and analyzed in terms of singular value decomposition modes. The structure and the interpretation of these modes is discussed. We use the measured STIR functions to compute the principal components of response to Walsh patterns, which are then compared with the results of McClurkin et al. (1991a,b). We also estimate the mutual information for our Walsh pattern Gedanken experiment and compare it with the same quantity measured by McClurkin et al. (1991a). In DISCUSSION we relate the STIR modes and principal components of the response to the issue of neural codes.

Preliminary aspects of this work have appeared (Shraiman et al. 1993).

METHODS

The data we use include data previously taken as part of a study on the chromatic properties of single units in the LGN of macaque monkey (Reid and Shapley 1992), as well as unpublished data. In total, we use the results for 9 magnocellular single units, 6 on-center and 3 off-center, and 31 parvocellular units. The parvocellular units are divided into subclasses that are based on the spectral sensitivity of their cone inputs, i.e., the short, medium, and long wavelength-sensitive cones denoted S, M, and L, respectively. There are three S on-center, one S off-center, five M on-center, four M off-center, nine L on-center and nine L off-center units. In two cases, both M off-center, the response of the unit is measured twice. This allows us to check the consistency of the data.

Data collection is as described (Reid and Shapley 1992). In brief, single tungsten electrodes are used to record from LGN relay cells in anesthesized and paralyzed macaque monkeys. The receptive field of magnocellular units lie between 3.0 and 23.0° of the fovea, and that of parvocellular units lie between 3.0 and 13.0°. A series of crossword puzzle–like patterns, constructed from m-sequences (Sutter 1987), are presented at fixed intervals. These patterns consist of an L_m by L_m matrix of squares that are chosen pseudorandomly to be either dark (labeled -1) or light (labeled $+1$). A sequence of patterns corresponds to a time-ordered list of -1's and $+1$'s for each of the L_m^2 squares. This sequence defines the stimulus, $S(\vec{r}, t)$. The spatial dimensions $\vec{r} = (x, y)$ are quantized in L_m steps, where $L_m = 8$ or 16 and each step subtends an angle of 0.13–0.43°. The temporal dimension, t, labels the pattern and is quantized in units of the stimulus frame interval, 14.8 ms. Only a tiny fraction of the $2^{L_m^2}$ patterns are shown in a given sequence,[1] whose length, N_m, is typically $N_m = 2^{16} - 1 = 65{,}535$. The contrast of the stimulus is 25% for 7 of the magnocellular units, 100% for 2 of these units, and 100% for all 31 parvocellular units.

The STIR function of a unit, denoted $R(\vec{r}, t)$, is found by correlating the measured spike train, $\Lambda(t)$, with the stimulus, i.e.

$$R(\vec{r}, t) = \frac{1}{T}\int_0^T dt' S(\vec{r}, t')\Lambda(t - t') \qquad (1)$$

In practice the RF is calculated for a finite interval of time, $t < 246$ ms, which is much shorter than the duration of the stimulus, $T = N_m \times 14.8$ ms. This interval corresponds to $N_T = 16$ frames, which is sufficient to record the STIR. For clarity in the formalism, all functions are written in terms of continuous variables, although they are treated as discrete during numerical calculations.

RESULTS

STIR function

LINEAR RESPONSE. The reverse correlation construction (*Eq. 1*) of the STIR function is founded on the assumption

[1] The value of the spatially averaged pair-wise correlations are $1/N_m$, as opposed, e.g., to $1/\sqrt{N_m}$ for sequences of patterns selected at random.

that the trial averaged activity of a cell is a function of a linear superposition of the inputs, i.e.

$$Z(t) = g\left[Z_0 + \int d^2r \int_{-\infty}^{t} dt' R(\vec{r}, t - t') S(\vec{r}, t') \right] \quad (2)$$

where the response of the cell is quantified by $Z(t)$, the probability of a spike being fired at time t or, equivalently, the instantaneous firing rate. The function $g(x)$ specifies the input-output relation, and the constant Z_0 controls the spontaneous firing level of the neuron. Provided that the stimulus dependent contribution is small compared with Z_0, so that stimulus induced modulation is small compared with the spontaneous firing rate,[2] the input-output relation (*Eq. 2*) can be linearized about $Z_0 = g(Z_0)$, i.e.

$$Z(t) = Z_0 + g'(Z_0) \int d^2r \int_{-\infty}^{t} dt' R(t - t', \vec{r}) S(\vec{r}, t') \quad (3)$$

The reconstruction of $R(t, \vec{r})$, up to a scale factor $g'(Z_0)$, via the reverse correlation *Eq. 1* for the m-sequence stimulus $S(\vec{r}, t')$, then follows from the assumption that the time average in *Eq. 1* is equivalent to the trial average for repetitions of the same spatial stimulus.

The STIR function, $R(\vec{r}, t)$, for two representative units are shown in Fig. 1; an on-center magnocellular unit and a long-wavelength–sensitive off-center parvocellular unit. The data are in the form of successive frames that are acquired at 14.8-ms intervals and quantized into 16 × 16 spatial pixels. Positive responses are coded green and negative responses red. We observe that for both units there is little discernible response until the third frame (t = 44 ms) and that the response peaks rapidly, by approximately the fourth frame (t = 59 ms). The well-described center-surround spatial structure, where the response at the center of the cell is opposite in sign from that in the surround, is evident in the magnocellular response but less clear in the parvocellular response. The spatial structure of other units is qualitatively similar. As time progresses, the sign of both the center and surround are seen to change.

DECOMPOSITION OF THE STIR FUNCTION. The resultant STIR functions are in general nonseparable functions of space and time, i.e., $R(\vec{r}, t) \neq F(\vec{r})G(t)$, where $F(\vec{r})$ is some function of space and $G(t)$ is some function of time. However, $R(\vec{r}, t)$ can be expressed as a sum of products of spatial and temporal modes, i.e.

$$R(\vec{r}, t) \equiv \sum_{n=1} \lambda_n F_n(\vec{r}) G_n(t) \quad (4)$$

where the spatial modes $F_n(\vec{r})$ and the temporal modes $G_n(t)$ from orthogonal bases, i.e.

$$\int d^2r F_n(\vec{r}) F_{n'}(\vec{r}) = \delta_{nn'} \quad \text{and} \quad \frac{1}{T} \int_0^T dt G_n(t) G_{n'}(t) = \delta_{nn'} \quad (5)$$

where $\delta_{nn'}$ is the Kronecker delta function. This expansion, formally known as a singular value decomposition (SVD), provides a simple description of the RF when few terms contribute to the sum (Golub and Kahan 1965). The calculation of the modes and expansion coefficients, λ_n, from the measured form of $R(\vec{r}, t)$ is described in APPENDIX A.

The representation of receptive fields in terms of the above decomposition is illustrated in Fig. 2 for the representative magnocellular and parvocellular units (Fig. 1). Only the first two terms are significant for the magnocellular unit. The first term consists of a symmetric unipolar spatial mode accompanied by a biphasic temporal mode, whereas the second term consists of a bimodal spatial mode accompanied by a triphasic temporal response. Interestingly, the center-surround structure appears in the minor mode. The representative parvocellular unit has three significant terms. As in the case of the magnocellular unit, the first term consists of a unipolar spatial mode. However, although the spatial structure of the high-order modes is bimodal, it is asymmetric and thus not described as center surround. In the above examples, and in general, the first term dominates. A statistical analysis (APPENDIX A) shows that, for 37 of the 40 units, at least 2 terms are significant. The ratio of the expansion coefficients is, on average, $|\lambda_1| : |\lambda_2| \simeq 4:1$. For five of the units, the third moment is significant only at the level of one standard deviation of the experimental noise level.

We now consider the form of the dominant temporal modes, $G_1(t)$ and $G_2(t)$, in detail (Fig. 3). The first-order mode for both magnocellular and parvocellular units peaks 45–60 ms after the onset of stimulation. The sign of the response then reverses, i.e., the response is bipolar, with the magnitude of the reversal particularly pronounced for magnocellular units. The response for both units recovers to the baseline value by 140 ms. The second-order mode is, not unexpectedly, more complex than the first-order mode. It appears triphasic for magnocellular units and biphasic for parvocellular units. Qualitatively, the temporal modes of both units are essentially the same for the first 60 ms, after which the parvocellular response decays considerably faster than that of magnocellular units. This later response is the origin of the descriptors "phasic" for magnocellular units and "tonic" for parvocellular units.

The above results show that the RF of units in the LGN are well approximated by the sum of only two space-time products. This suggests that a useful measure of the nonseparability between space and time is the normalized value of the coefficient for the second mode, i.e., $|\lambda_2|/(|\lambda_1| + |\lambda_2|)$. The values of this measure are broadly distributed, with a mean of 0.22 (Fig. 4). There are no apparent differences between magnocellular and parvocellular units.

Comparison with the measurements of McClurkin et al.

PREDICTED RESPONSE TO WALSH PATTERNS AND THE PRINCIPAL COMPONENTS. We consider first the relation between the STIR functions reported in this work (Figs. 2 and 3) and the results of McClurkin et al. (1991a) on the response of units to Walsh patterns. Like the m-sequence patterns, Walsh patterns consist of black and white squares (e.g., Fig. 4). Each pattern has L_W squares on edge, or L_W^2 squares total. They form a complete basis, in the sense that linear combinations of different patterns can represent any black

[2] The modulation amplitude of the response is expected to be small in early visual areas for stimuli with sufficiently low contrast. The integrated stimulus contrast in the present experiments was observed to maintain the output of most magnocellular and all parvocellular units in their linear range.

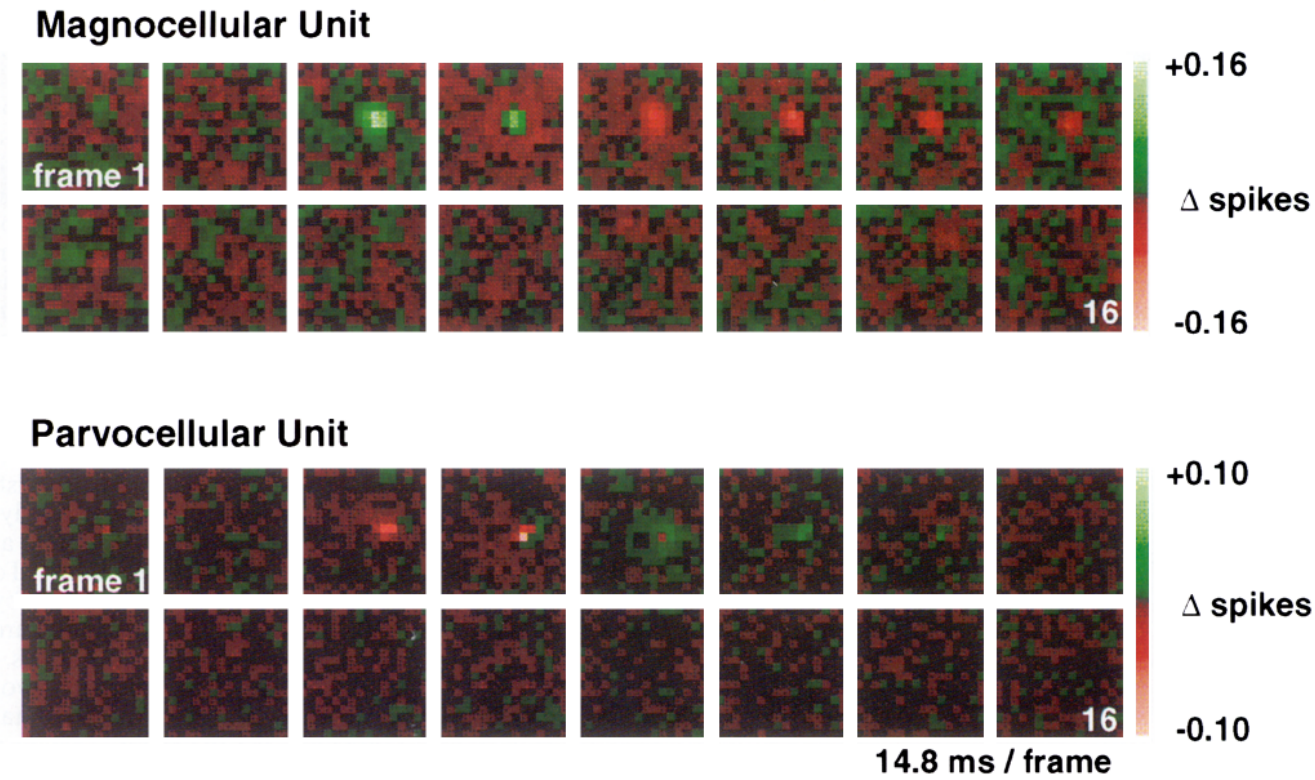

FIG. 1. Space-time receptive field (RF) for representative units in the lateral geniculate nucleus (LGN). Space in quantized in pixels, with 16 × 16 pixels per frame, and time is quantized in frames, corresponding to the 14.8-ms refresh period of the m-sequence stimuli. Changes in firing rate are color-coded, as indicated. *A*: results for an on-center magnocellular unit [*uq021101.tin*]. Each pixel is 0.43° on edge. The scale is in spikes/frames above the background level of 0.52 spikes/frame, or 33.7 spikes/s, and is chosen to highlight the average activity; the largest observed change in any pixel is 0.16 spikes/frame. *B*: results for a long-wavelength off-center parvocellular unit [*uq051304.tin*]. Each pixel equals 0.13° on edge. The background level is 0.53 spikes/frame, or 34.5 spikes/s; the largest observed change is 0.10 spikes/frame.

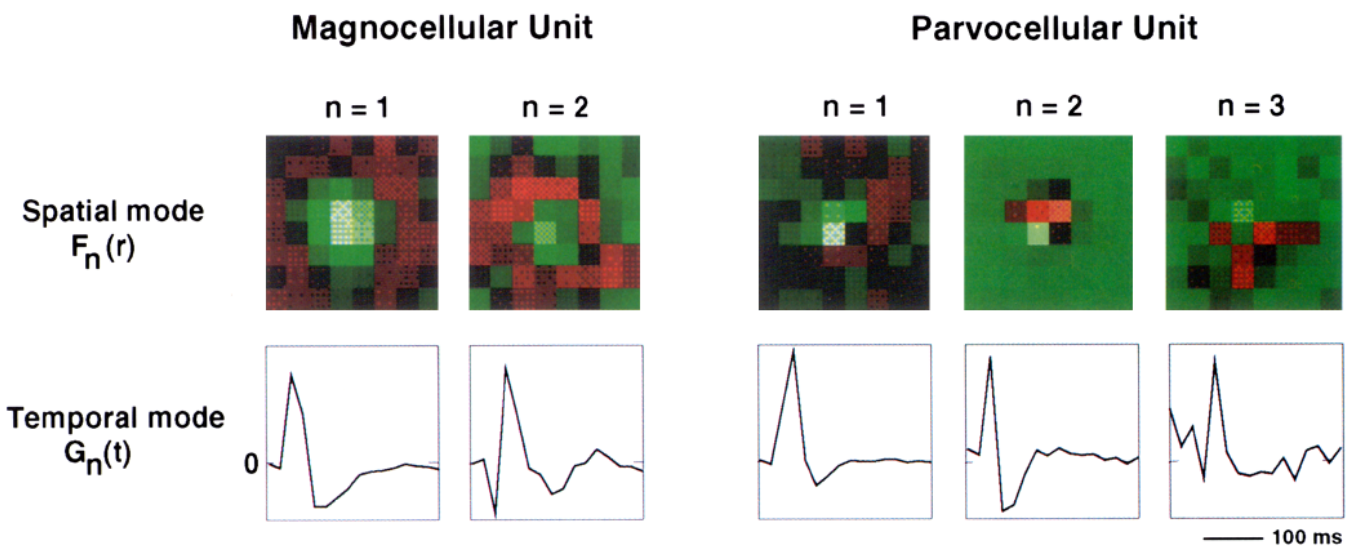

FIG. 2. Singular value decomposition of the RF for LGN units. Shown are the spatial modes $F_n(\vec{r})$ and the temporal modes $G_n(t)$ (*Eq. 4*) for the representative units in Fig. 1. The spatial modes include only the 8 × 8–pixel subregion containing the active part of the field. They are presented as false colored images with red indicating positive values and green indicating hyperpolarization. *A*: results for the 1st 2 modes of the on-center magnocellular unit. The expansion coefficients are $\lambda_1 = 0.840\ s^{-1}$, $\lambda_2 = -0.233\ s^{-1}$, $\lambda_3 = -0.125\ s^{-1}$, and $\lambda_4 = 0.119\ s^{-1}$; only the 1st and 2nd terms in the expansion are statistically significant. Note that only the *ratio* between the absolute values of the eigenvalues is meaningful. *B*: results for the 1st 3 modes of the long-wavelength off-center parvocellular unit. The expansion coefficients are $\lambda_1 = 0.634\ s^{-1}$, $\lambda_2 = -0.229\ s^{-1}$, $\lambda_3 = -0.064\ s^{-1}$ and $\lambda_4 = -0.046\ s^{-1}$. The 1st 3 terms in the expansion are statistically significant. Note that the magnitude of λ_2 for this particular unit is atypically large.

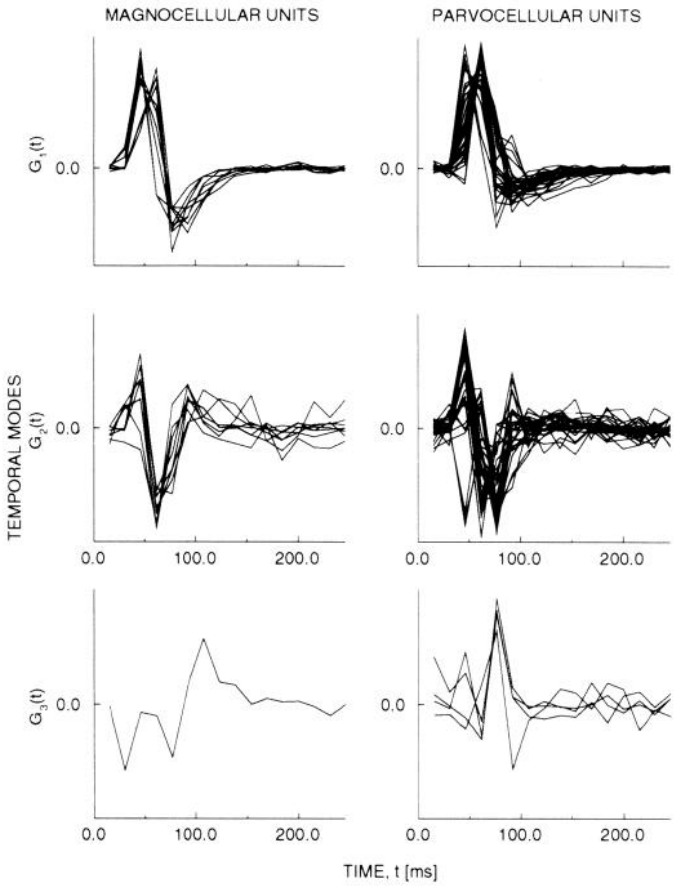

FIG. 3. Dominant temporal modes, $G_1(t)$, $G_2(t)$, and $G_3(t)$, for our magnocellular and parvocellular units. For $G_2(t)$ and $G_3(t)$ we show the waveform only for those units in which the expansion coefficients are statistically significant.

and white picture with a resolution of 1 part in L_W. For this case, the stimulus is of the form

$$S_\alpha(\vec{r}, t) = \begin{cases} u_\alpha(\vec{r}) & \text{if } 0 < t < T \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $u_\alpha(\vec{r})$ defines the spatial pattern of the αth stimulus and includes both normal and contrast reversed images. These patterns satisfy

$$\frac{1}{N_W} \sum_{\alpha=1}^{N_W} u_\alpha(\vec{r}) u_\alpha(\vec{r}') = \delta_{rr'} \tag{7}$$

where $N_W = 2L_W^2$ is the number of patterns. The sum of all patterns is a blank, i.e.

$$\frac{1}{N_W} \sum_{\alpha=1}^{N_W} u_\alpha(\vec{r}) = 0 \tag{8}$$

Using *Eq. 2* we obtain $Z_\alpha(t)$, the average neuronal spiking activity at time t after the onset of Walsh pattern α, i.e.

$$Z_\alpha(t) = g\left[Z_0 + \int dr^2 \int_0^{\min(t,T)} dt' R(\vec{r}, t - t') u_\alpha(\vec{r}) \right] \tag{9}$$

As emphasized earlier (*Eq. 3*), for sufficiently weak stimuli *Eq. 9* can be linearized and the stimulus-induced variation in Z_α is determined by $R(\vec{r}, t)$ up to a multiplicative constant. For stronger stimuli an additional assumption about the form of $g(x)$ is needed. A minimal such assumption is

$$g(x) = \begin{cases} x & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases} \tag{10}$$

which corresponds to rectification that prevents the instantaneous firing rate $Z(t)$ from having non-negative values. The rectification effect is important only when negative modulation induced by stimuli are comparable with the spontaneous firing rate. We do not include the effect of saturation in *Eq. 10* on the assumption that the maximum firing rate of LGN neurons, on the order of 100 spikes/s, is never reached in the experiments that we consider. *Equations 9* and *10*, combined with the measured STIR functions as parameterized by *Eq. 4* and estimates of the background firing rate, Z_0, and stimulus amplitude, $|u_\alpha|$, allow one to compute the expected temporal response to the Walsh patterns. An example for a particular parvocellular unit is shown in Fig. 5, where we used a 4×4 set of Walsh patterns and include only one sign of contrast. The steady-state change in firing rate as well as the transient change at short times is seen to vary significantly between stimuli. To compare the predicted response with those reported by McClurkin et al. (1991a,b), we need to consider a measure of the ensemble averaged response of parvocellular units.

McClurkin et al. (1991b) measure the response of parvocellular units in the LGN averaged over several presentations for each stimuli comprising the Walsh set. They find that each of the measured responses, $Z_\alpha(t)$, is accurately represented by a small number of temporal modes, denoted the principal components[3] $\Phi_n(t)$, and express their results in the form

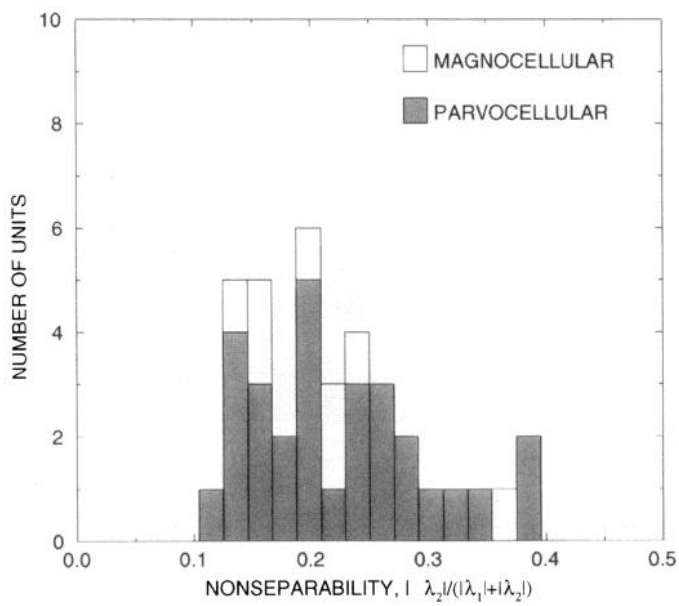

FIG. 4. Quantification of nonseparability for units in the LGN. Shown is a histogram of the nonseparability between space and time for 37 of the 41 units in which at least 2 terms in the singular value decomposition (*Eq. 4*) are significant. Open regions correspond to magnocellular units, and shaded regions to parvocellular units.

[3] In the present work the principal components are labeled $\Phi_1(t), \ldots$, whereas in the work of McClurkin et al. the indexing starts at 0 and the components are $\phi_0(t), \ldots$.

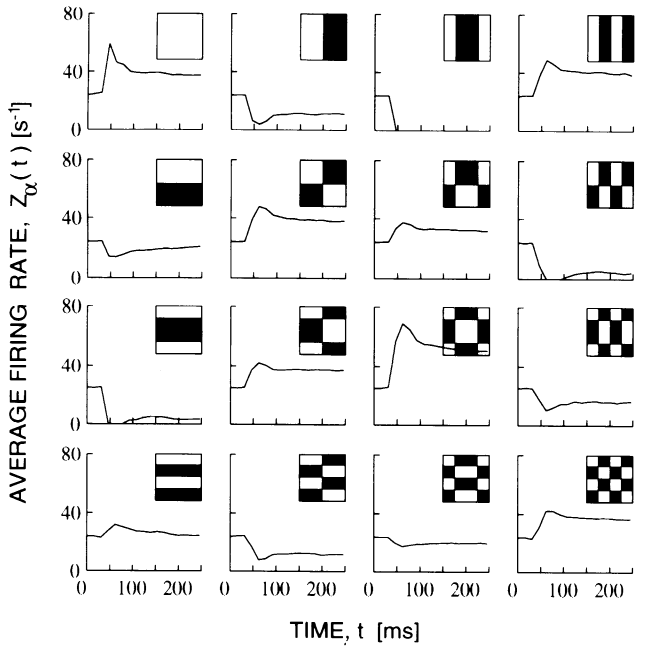

FIG. 5. Average temporal response, $Z_\alpha(t)$, calculated (*Eqs. 9* and *10*) for all members of a 4 × 4 set of Walsh patterns with the use of our representative parvocellular unit (Figs. 1*B* and 2*B*). *Inserts* show the particular pattern.

$$Z_\alpha(t) = \bar{Z}(t) + \sum_{n=1} a_{\alpha,n}\Phi_n(t) \qquad (11)$$

where $\bar{Z}(t)$ is the average response to all of the stimuli, i.e.

$$\bar{Z}(t) = \frac{1}{N_W}\sum_{\alpha=1}^{N_W} Z_\alpha(t) \qquad (12)$$

The principal components are by definition the eigenfunctions of the covariance matrix, $C(t, t')$, of the measured averaged neuronal responses, i.e.

$$C(t, t') = \frac{1}{N_W}\sum_{\alpha=1}^{N_W} [Z_\alpha(t) - \bar{Z}(t)][Z_\alpha(t') - \bar{Z}(t')] \qquad (13)$$

It is important to stress that the responses in the covariance matrix are already averaged over all trials, i.e., repetitions of a given stimulus. Thus the covariance matrix defined above does not include trial-to-trial fluctuations. Last, the expansion coefficients $a_{\alpha,n}$ are

$$a_{\alpha,n} = \frac{1}{T}\int_0^T dt Z_\alpha(t)\Phi_n(t) \qquad (14)$$

To make contact with McClurkin et al. (1991a), we perform a detailed calculation of the principal components for each of our magnocellular and parvocellular units (*Eqs. 4* and *6–13*). The first three principal components are shown in Fig. 6. The transient behavior of $\Phi_1(t)$ and $\Phi_2(t)$ is confined to early times, as expected from the decomposition of the RF (Fig. 2). There is a spectrum of waveforms for the principal components calculated for the magnocellular units (Fig. 6, *A–C*). For several magnocellular cells, the first principal component eigenvector does approach the baseline at long times. The reason for this diversity is unknown. On the other hand, the form of the principal components is quite similar for all parvocellular units (Fig. 6, *D–F*).

The principal components we predict for parvocellular units on the basis of the measured RFs (Fig. 6, *D–F*) compare well with those reported by McClurkin et al. (1991a) (reproduced in Fig. 6, *G–I*). The shape and time course of the predicted and measured forms of $\Phi_1(t)$ and $\Phi_2(t)$ are, qualitatively, indistinguishable at short times. There is a small, slow component in the second component reported by McClurkin et al. (1991a) that is not present in our results. This is likely to be a consequence of adaptation during their relatively long period of stimulation (see DISCUSSION). The third principal component in the analysis of McClurkin et al. (1991a) is essentially insignificant, similar to the predicted result. Two of the units that comprise the data of McClurkin et al. (1991a) are reported to be atypical (dashed and dotted lines in Fig. 6, *G–I*). We suggest that at least one of these units is a magnocellular unit (cf. dashed line in Fig. 6, *G–I*, with Fig. 6, *A–C*).

QUANTITATIVE MEASURES OF STIMULUS DISCRIMINATION. McClurkin et al. (1991a) have observed that the inclusion of time dependence in the measures of neuronal response enhances the ability to discriminate between the distinct stimuli. A quantification of discrimination is the mutual information between the set of stimuli and the response, and an increase in mutual information is consistent with the general notion (e.g., Cover and Thomas 1991) that the mutual information between a fixed set of inputs and a set of outputs can only increase with an increase of the dimensionality of the output space. In other words, the mutual information between the stimuli and the set of measurements of the neuronal response will only increase as additional measurements of the response are made. For example, the output space is one dimensional if only the total number of spikes in the measurement period is reported. It is two dimensional if one measures projections onto two principal components, and it is K dimensional if the response is described by the instantaneous firing rate measured at K points in time. Note that the gain in the mutual information occurs only to the extent that different measurements are not completely correlated with each other while still correlated with the stimulus. This requirement makes the principal components of the response a sensible choice of basis, as we shall explain in the following section.

We now estimate quantitatively the expected gain in mutual information due to the increase in temporal resolution of the response. The estimate for parvocellular units will be directly compared with the results of McClurkin et al. (1991a). To compute the mutual information between the spike train $\Lambda(t)$ observed in a single trial and the stimulus, one needs to know the statistics of the spike train in addition to its average instantaneous firing rate (*Eq. 2*). We shall assume the spikes to be generated by an inhomogeneous Poisson process (APPENDIX B).

Tables 1 and 2 show the mutual information calculated for three characterizations of the response with increasing complexity (APPENDIX C): *1*) the total number of spikes, $\bar{\Lambda}$ (*Eq. C5*); *2*) the overlap of the spike train with the first principal component, $\bar{\Lambda}_1$ (*Eq. C4*); and *3*) the complete spike train, $\Lambda(t)$ (*Eq. B6*). For these calculations the presentation time is fixed at 246 ms, close to the value of 256 ms

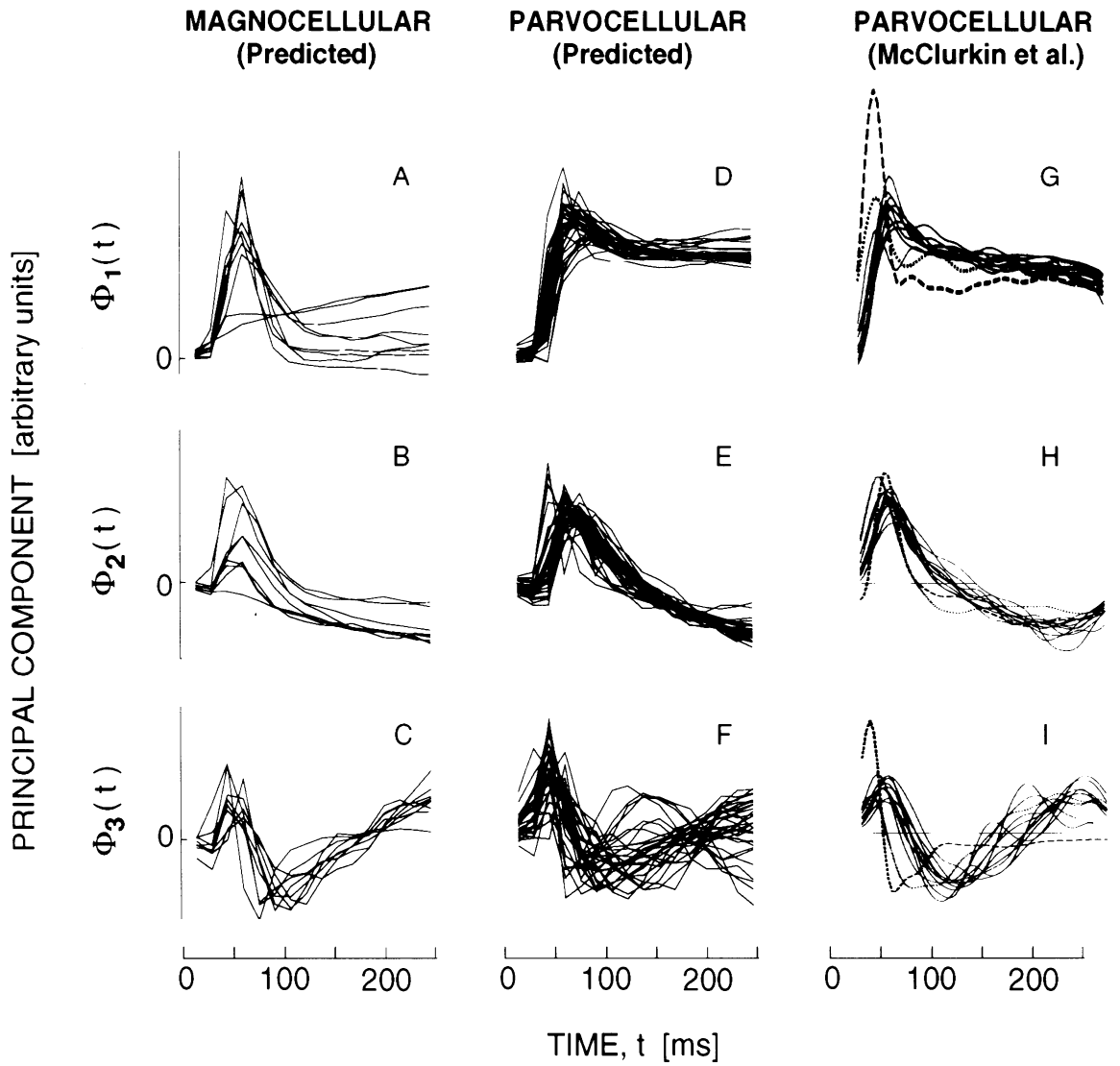

FIG. 6. Principal components of the neuronal output in response to Walsh pattern stimuli. The functions $\Phi_1(t)$, $\Phi_2(t)$, and $\Phi_3(t)$ are calculated for all of our units, as described (*Eq. 6*), and are compared with those reported by McClurkin et al. (1991a). *A–C*: results for the magnocellular units. *D–F*: results for our parvocellular units. *G–I*: principal components reported by McClurkin et al. (1991a) from measurements on parvocellular units in the LGN. The 2 dashed lines correspond to units that are judged by those authors to be atypical. These data should be contrast with the components calculated for our parvocellular units, cf. *D* and *G*, *E* and *H*, and *F* and *I*.

used in the experiments of McClurkin et al. (1991b). We observe a doubling of the mutual information for our magnocellular units in comparing the response for the full spike train versus the number of spikes (Table 1) but only a 30% increase for parvocellular units (Table 2). The greater increase in mutual information for magnocellular units reflects the transient nature of their response characteristics (Fig. 3), an issue we explore by considering the dependence of the mutual information for the three above cases on the presentation time of the stimulus (Fig. 7).

TABLE 1. *Mutual information for different measures of neuronal response: magnocellular units*

Measure	Predicted Value
$I(\bar{\Lambda}; S)$ Number of spikes	0.28 ± 0.05
$I(\bar{\Lambda}_1; S)$ Overlap of train with $\Phi_1(t)$	0.41 ± 0.04
$I(\Lambda; S)$ Spike train	0.60 ± 0.06

Values in Predicted Value are means ± SE; number of units in Predicted Value is 9.

Information based on the number of spikes, $I(\bar{\Lambda}; S)$. We focus first on our representative magnocellular unit (Fig. 1*A*). The mutual information rises steeply from chance, $I(\bar{\Lambda}; S) = 0$, at short integration times; achieves a maximum value as the time increases; and then decays slightly to a steady-state plateau value at long times (triangles; Fig. 7*A*). The initial rise occurs because the integrated activity for magnocellular units is greatest during the early part of the response [see $\Phi_1(t)$ in Fig. 7*A*]. The slight dip and plateau occur because the integrated response receives relatively little contribution from stimulus related events

after the first 50 ms but continued contributions from background firing.[4] In contrast to the case for the magnocellular unit, the mutual information for the representative parvocellular unit rises essentially continuously over the entire time course of stimulation (triangles; Fig. 7*B*). This behavior is a consequence of the sustained response of parvocellular units at long times.

Information based on the first principal component, $I(\bar{\Lambda}_1; S)$. The first principal component provides the dominant contribution to the average response of our units (Fig. 4) and, as shown later, dominates the reliability of parvocellular units (Fig. 8). We observe a significant increase in the estimate of mutual information based on the overlap of the spike train with the first principal component compared with the information calculated for the number of spikes (cf. squares with triangles in Fig. 7, *A* and *B*). The increase is 32% for magnocellular units and 12% for parvocellular units. The greater increase for magnocellular units reflects the relatively limited time interval over which they respond.

Information based on the complete spike train, $I[\Lambda(t); S]$. For this case the mutual information must be a monotonically increasing function of time. We observe that, for the magnocellular unit, the mutual information shows a sustained albeit small rise at long times in addition to the rapid rise at short times discussed above (circles; Fig. 7*A*). The latter rise reflects a relatively small but nonetheless significant steady-state component in the response of this unit. For the case of the parvocellular unit, the time course of the mutual information behaves quite similar to that calculated for the reduced measures (cf. circles with squares and triangles in Fig. 7*B*). This occurs because the integrated value of the temporal response for parvocellular maintains a significant plateau (Fig. 3) with no discernible feature(s).

COMPARISON WITH THE MEASUREMENTS OF McCLURKIN ET AL. We compare our predicted results of the mutual information for parvocellular units with that found in the experiments of McClurkin et al. (1991a). Within uncertainty, the mutual information between the full spike train and the Walsh patterns is the same in both studies, ~0.7 bits (Table 2). Further, when the number of spikes is considered, rather than the full train, a reduction of the mutual information by 20–30% is seen from both studies. McClurkin et al. (1991a; Optican et al. 1991) report that the mutual information is significantly reduced when only the overlap of the spike train with the first principal component is considered. In contrast, we predict a smaller effect (Table 2). The overall agreement between the two studies is surprisingly good, perhaps better than one has a right to expect in view of difference in the experimental conditions between our measurements and those of McClurkin et al. (1991b) and in light of the assumptions made in our analysis (APPENDIX C and DISCUSSION).

[4] A similar conclusion is reached by Tovée et al. (1993) for the information content of units with phasic response properties in primate temporal visual cortex.

TABLE 2. *Mutual information for different measures of neuronal response: parvocellular units*

Measure	Predicted (Present Study)	McClurkin et al. (1991a)
$I(\bar{\Lambda}; S)$ Number of spikes	0.59 ± 0.04	0.47 ± 0.06
$I(\bar{\Lambda}_1; S)$ Overlap of train with $\Phi_1(t)$	0.67 ± 0.05	0.48 ± 0.07
$I(\Lambda; S)$ Spike train	0.75 ± 0.05	0.64 ± 0.10

Values in Predicted and McClurkin et al. are means ± SE; number of units in Predicted is 31 and in McClurkin et al. is 11.

Principal components and coding

We now discuss the meaning of the modes found by our singular value decomposition of the spatiotemporal response function as well as their relation to the principal components of the response to Walsh patterns measured by McClurkin et al. (1991a) and to issues of information and coding raised by these authors (Gawne et al. 1991; McClurkin et al. 1991c).

The dominant SVD modes describe those aspects of the stimulus that control the instantaneous, trial-averaged firing rate of the unit at a given poststimulus time. Thus, for example, the spatial patterns of a stimulus orthogonal to the first two (or 3) spatial modes do not contribute to the response, i.e., the unit is "blind" to those aspects of the stimulus. Also, because the spatial modes are orthogonal to each other, they correspond to different "features" of the stimulus and thus are, in principle, independent. To the extent that these independent features are discernible in the output, one can speak of them as being "encoded" in the output.

In general, the instantaneous response depends not only on the spatial but also on the temporal aspects of the stimulus. For the special case that the time dependence of the stimulus is particularly simple, e.g., a stationary stimulus during a fixed presentation time T, a simple relation between orthogonal spatial modes of the stimulus and orthogonal temporal modes of the response emerges. In the *linear* regime, these temporal modes are found by the SVD analysis of response to pulses of duration T, obtained by using $S(\vec{r}, t) = u(\vec{r})$ for $0 < t < T$ in *Eq. 2*, i.e.

$$Z(t) = g\left[Z_0 + \int d^2r R_T(\vec{r}, t)u(\vec{r})\right] \tag{15}$$

where, as before, $u(\vec{r})$ refers to the spatial pattern of the stimulus and

$$R_T(\vec{r}, t) \equiv \int_0^{\min(t,T)} dt' R(\vec{r}, t - t') \tag{16}$$

The SVD analysis of $R_T(\vec{r}, t)$ (APPENDIX A) generates a set of orthogonal spatial and temporal modes, $\tilde{F}_{n;T}(\vec{r})$ and $\tilde{G}_{n;T}(t)$, analogous to those we found for $R(\vec{r}, t)$ (*Eq. 4*).[5] On the other hand, when *Eq. 15* can be linearized (*Eq. 3*) the $\tilde{G}_{n;T}(t)$ are, by their definition, the principal components of trial-averaged responses $Z_\alpha(t)$ for a complete set of stimuli $S_\alpha(\vec{r}, t)$ (*Eqs. 6–8*). Aside from a constant factor,

[5] The SVD modes of $R_T(\vec{r}, t)$ reduce to those of $R(\vec{r}, t)$ in the limit $T \to 0$.

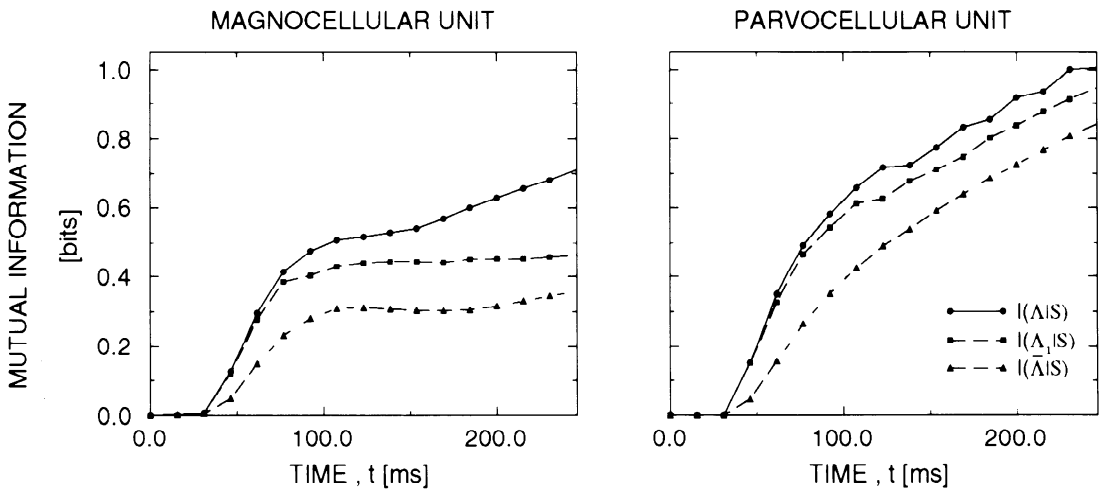

FIG. 7. Reliability of representative units for discriminating between stimuli on the basis of the neuronal output. *A*: mutual information between the neuronal output and the stimulus (*Eq. C2*) for the representative magnocellular unit (Figs. 1*A* and 2*A*). The solid curve with circles is the measure for the full spike train, $I(\Lambda; S)$, whereas the dashed line with squares is for the first principal component, $I(\Lambda_1; S)$, and the one with triangles is for the number of spikes, $I(\bar{\Lambda}; S)$. *B*: mutual information between the neuronal output and the stimulus for the representative parvocellular unit (Figs. 1*B* and 2*B*).

they differ from the $\Phi_n(t)$ found in the previous section (*Eq. 11*) only to the extent that $Z_\alpha(t)$ computed for the set of Walsh patterns is affected by rectification (*Eq. 10*). Thus, for the case of pulse stimuli, the spatial mode $\tilde{F}_{n;T}(\vec{r})$ is encoded in the temporal response as a principal component $\tilde{G}_{n;T}(t)$. This establishes the relation between the SVD of the response function, the principal components, and the notion of coding as proposed by Gawne et al. (1991).

An alternative point of view is that the principal components should be considered as independent information channels. To make this notion precise, consider the additional discrimination capability provided by the inclusion of an additional SVD mode or principal component in the measured "output" of a neuron. The amount depends on the magnitude of the contribution to the response made by this mode compared with the root-mean-square (RMS) fluctuations of the response. To illustrate this point we again consider neuronal responses in the linearized regime so that *Eq. 15* becomes

$$Z(t) = Z_0 + \sum_n \lambda_n A_n \tilde{G}_{n;T}(t) \tag{17}$$

with the spatial structure of the stimulus parameterized by the projections

$$A_n \equiv \int d^2r \tilde{F}_{n;T}(\vec{r}) u(\vec{r}) \tag{18}$$

If $Z(t)$ were known exactly, all the stimulus parameters would be determined exactly as well. The question, however, is how well the parameters can be estimated without the precise knowledge of $Z(t)$, e.g., from a single spike train $\Lambda(t)$ of duration T that is generated by an inhomogeneous Poisson process with instantaneous rate $Z(t)$. The simplest estimate of A_n is

$$\hat{A}_n = \frac{1}{\lambda_n T} \int_0^T dt \tilde{G}_{n;T}(t)[\Lambda(t) - Z_0] \tag{19}$$

The trial average of the estimator is $\langle \hat{A}_n \rangle_{\text{trial}} = A_n$, found from *Eq. 19* with $\langle \Lambda(t) \rangle_{\text{trial}} = Z(t)$ and *Eqs. 5* and *18* and strictly valid in the limit of an infinite number of trials. Note that the size of A_n depends on the change in firing induced by the stimulus as well as on the SVD modes for the unit. An estimate of the maximum size of A_1 for our parvocellular units, for which $\lambda_1 \sim 1\ \text{s}^{-1}$ because the first mode dominates and for which $\tilde{G}_{1;T}(t)$ is approximately a constant [$\tilde{G}_{1;T}(t) \simeq \phi_1(t)$ in Fig. 6*D*], is $|A_1| < (Z_0 T)/(\lambda_1 T) \sim 10$.

To assess the expected RMS fluctuations of the estimator, we consider the variance of $\hat{A}_n$ for a Poisson spike process, or, more properly, the covariance σ^2_{nm} of the estimators $\hat{A}_n$ and $\hat{A}_m$ (APPENDIX D), i.e.

$$\sigma^2_{nm} = \langle (\hat{A}_n - A_n)(\hat{A}_m - A_m) \rangle_{\text{trial}} \simeq \frac{Z_0 T}{(\lambda_n T)^2} \delta_{nm} \tag{20}$$

The form of *Eq. 20* shows that different $\hat{A}_n$ are uncorrelated and that the RMS fluctuations are inversely proportional to the eigenvalue λ_n associated with the SVD mode, so that modes with small λ_n are difficult to estimate precisely. The scale of the RMS fluctuations is set by $\sqrt{Z_0 T}$, where $Z_0 T$ is just the average number of background spikes during the observation of the response.[6] For our parvocellular units, the RMS fluctuations are $\sqrt{\sigma^2_{11}} \sim 3$ (see above), and the magnitude of the stimulus parameter A_1 is at most approximately three times the level of fluctuations in the estimate of A_1 based on a single trial, i.e., $|A_1|/\sqrt{\sigma^2_{11}} < 3$.

The present calculation suggests that the different temporal modes, or principal components, can be viewed as independent information channels with higher order channels becoming increasingly unreliable. Further, it allows us to illustrate why the addition of the second channel, i.e., the inclusion of the additional principal components in the characterization of the response, does not re-

[6] If the response $\tilde{G}_{n;T}(t)$ decays sufficiently rapidly with time to be square integrable, the T^{-1} normalization factor in *Eqs. 19* and *20*, as well as *Eq. 5*, can be omitted. With this normalization, σ^2_{mn} does not depend on the observation time, T, as would be the case for the magnocellular response.

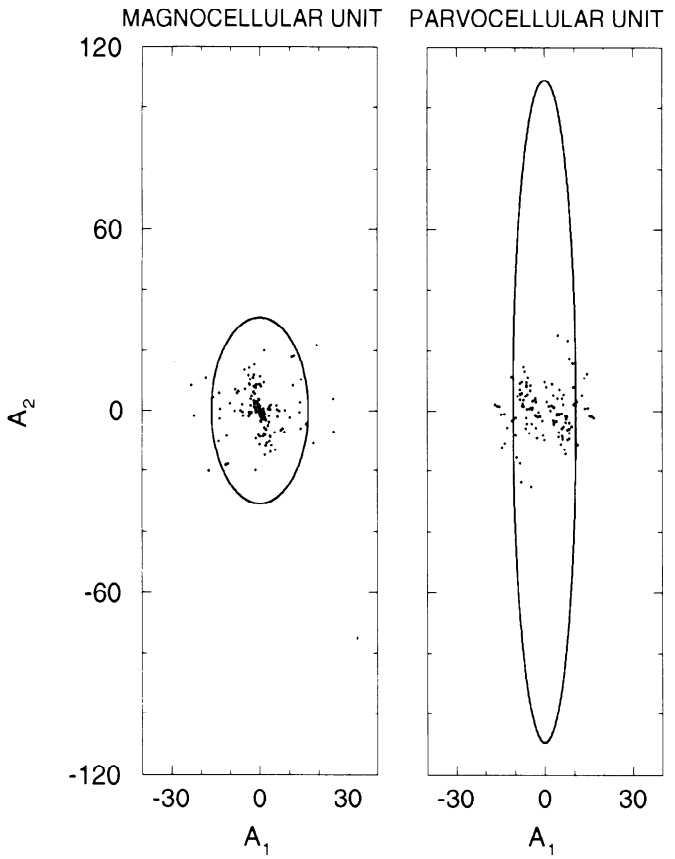

FIG. 8. Root-mean-square fluctuations of the neuronal response based on a single trial. Shown are calculations for the representative magnocellular and parvocellular units (Figs. 1*A* and 2*A*). Average firing rates $Z_\alpha(t)$ have been reduced by 55% to ensure that the units operate in the linear regime; this corresponds to a reduction in contrast. Ellipses mark the 1 standard deviation boundary in the space of estimation parameters A_1 and A_2 with the use of the full spike train, i.e., $(A_1/\sigma_{11})^2 + (A_2/\sigma_{22})^2 = 1$ (*Eq. 20*). Superimposed on each figure is a scatter diagram of the projections for the trial-average response of the units to the 128 Walsh patterns (*Eqs. 18* and *19*).

sult in a large increase in the mutual information. We plot (Fig. 8) the distribution of the projections for all 128 Walsh stimuli in the (A_1, A_2) coordinate plane, calculated for our representative magnocellular and parvocellular units (*Eq. 18*), along with the ellipse whose minor and major axes correspond to the RMS fluctuations in the estimation of A_1 and A_2 from a single trial response, i.e., $\sqrt{\sigma_{11}^2}$ and $\sqrt{\sigma_{22}^2}$, respectively (*Eq. 20*). The estimation error represented by the ellipse is seen to be large compared with the relative spread in the values of the parameters A_1 and A_2 for different stimuli, i.e., each ellipse encloses the majority of the points in Fig. 8. This analysis explains the relatively poor discrimination performance of a single neuron, as suggested above. Further, although estimation of both A_1 and A_2 contribute significantly to the discrimination capabilities of the unit, the fluctuations associated with the estimation of A_2 are relatively large for the parvocellular unit and result in an ellipse with a particularly elongated axis along A_2, i.e., $\sqrt{\sigma_{22}^2} \gg \sqrt{\sigma_{11}^2}$ (Fig. 8). This explains why there is little difference between the mutual information calculated based only on the first mode versus that based on the complete spike train for parvocellular units (Table 2).

DISCUSSION

We use our measurements of the RFs of parvocellular units to predict the average temporal response of these units to a set of Walsh patterns, as well as to predict the reliability of these units for distinguishing between patterns on the basis of a single response. These predictions are compared with the results of McClurkin et al. (1991a). Although our predictions provide a vehicle to demonstrate the possibility of such comparison, they are necessarily imprecise because of differences in the experimental conditions present in the two studies. The measurements reported here are performed on macaque monkeys that are anesthetized and mechanically respired. Those of McClurkin et al. (1991b) involve awake rhesus monkeys. In both studies one pixel in the stimulus encompasses slightly less than the central region of the RF, but detailed comparisons are impossible.

The emphasis in this work is on the temporal properties of units, and thus our data are taken under conditions that maximize temporal resolution at the expense of spatial resolution (Fig. 1). Nonetheless, there are features that can be discerned from the spatial modes of the RFs. First, we observe that the dominant contribution to both magnocellular and parvocellular units has a symmetric unipolar shape (Fig. 2). Thus objects with a circular shape are optimal stimuli for these LGN units. Second, the center-surround structure is present only in the secondary mode (Fig. 2) for magnocellular units and generally is not apparent in the second or higher order modes of parvocellular units, although the relatively low ratio of signal-to-noise in the data for the latter units (Fig. 1*B*) leads to a poor estimate of their spatial structure.

The functional form of the average responses is expressed in terms of a small number of temporal modes, known as principal components (*Eq. 13*). We observe qualitatively good agreement between the predicted modes and those reported by McClurkin et al. (1991a) (Fig. 6). A possible significant difference between the two measurements occurs only for the second mode at long times. This may be related to adaptation. The third components are marginally significant in both studies. With regard to the contrast of the stimuli, we find that a change in the ratio between the background rate and the stimulus-related modulation by a factor of two (in both directions) does not appreciably change the shape of the principal components (APPENDIX C).

The reliability of units in discriminating between different patterns is quantified in terms of mutual information. We observe good although not precise agreement between the predicted values and those reported by McClurkin et al. (1991a) (Table 2). Discrepancies between the two sets of values may arise from a number of sources. One is the difference in experimental conditions, as mentioned above. A second source of discrepancy may involve the assumptions that we use. The linear-threshold approximation for a neuron (*Eq. 10*) is not exact. Further, the statistics of LGN units deviate from those of an inhomogeneous Poisson pro-

cess, e.g., the neuronal refractory period causes the statistics to be non-Poisson shortly after a spike. We note, however, that Geisler et al. (1991) shows that the measured deviation from Poisson statistics for units in auditory nerve and visual cortex essentially does not affect their reliability. A final, possible source of discrepancy relates to the method used by McClurkin et al. (1991a) to calculate the mutual information from their measured responses (Optican et al. 1991). These workers smooth their spike trains with a Gaussian filter. The width of this filter affects the estimate of the mutual information (Optican et al. 1991). Recent methods introduced by these workers may alleviate this problem (Chee-Orts and Optican 1993; Hertz et al. 1992).

Quantifying the reliability of neurons

We focused on Shannon's mutual information as a measure of performance for discrimination tasks solely as a means to compare our results with those of McClurkin et al. (1991a). Although this measure is well defined (*Eq. C2*) and is used to characterize a number of sensory systems (e.g., Bialek et al. 1991), its interpretation in the context of discrimination tasks is problematic (Geisler et al. 1991). A different and possibly more natural measure of neuronal reliability is the probability of correct response (Geisler et al. 1991; Miller et al. 1993). This indicator reports the fraction of instances in which the stimulus is correctly identified from a single spike train. Its calculation depends on relating the best estimate of a stimulus, based on the observed spike train, to the stimulus itself. With respect to our parameterization of visual stimuli in terms of their projections on the spatial modes of the unit response (*Eq. 18*), the probability of correct response measures the area covered by the projections of all of the stimuli (dots in Fig. 8) relative to the area of the RMS fluctuations in the projections (ellipse in Fig. 8). Thus widely dispersed stimuli lead to high reliability, and vice versa.

Optimum rate of background firing

The reliability with which stimuli can be identified on the basis of a single spike train depends on the background rate (*Eq. 20* and APPENDIX B). When this rate is too low, there is a tendency for many stimuli to make the output of the neuron quiescent. This leads to poor reliability. On the other hand, when this rate is too high, the random nature of the spike train contributes excessive noise, and, again, the reliability is poor. There is thus an optimal background rate, whose value depends on average modulation of the spike train by natural stimuli. Interestingly, in our analysis of the response of units to Walsh patterns, we find that the background rates are typically within 50% of the optimal rate.

Is there a "neural code" for output from the LGN?

We demonstrate that the temporal structure of the neural response observed in the experiments of McClurkin et al. is consistent with that expected on the basis of our spatiotemporal RF data. The above authors motivate their study of the principal components of the response by notions of coding, i.e., the set of temporal principal components is interpreted as a finite set of "codebook" vectors that represent particular components of the spatial structure of the stimulus (Gawne et al. 1991). Indeed, such an interpretation appears natural in the context of a general linear mapping of a stimulus vector, S_a, into a response vector, Z_α, i.e., $Z_\alpha = \sum_a G(\alpha|a)S_a$. The singular value decomposition (*Eq. 4*) provides a representation for the map, $G(\alpha|a) \equiv \sum_n \phi_n(\alpha)\chi_n(a)$, so that the orthogonal response modes $\phi_n(\alpha)$ appear to code for the orthogonal input features $\chi_n(a)$. It is appealing to interpret this apparent relation in the context of experiment by identifying the input label "a" as a spatial coordinate of the stimulus and "α" as the time variable t of the response. However, the stimulus is itself time dependent and thus contributes to the time dependence of the output. Thus, in general, we must take $a = (\vec{r}, t')$ and identify $G(\alpha|a)$ as $R(t|\vec{r}, t') = R(\vec{r}, t - t')$ (*Eq. 2*). The SVD of $R(\vec{r}, t - t')$ yields a continuous spectrum of eigenmodes[7] and does *not* provide a finite set of principal component vectors that encode the stimulus. This is the consequence of the continuous temporal evolution of the response to a time-dependent stimulus. A finite set of principal components is obtained only in response to a stimulus of fixed duration and depends in an essential manner on the particular time course of the stimulus. Consequently, such principal components do not form a unique representation of the spatial features of the stimulus. Rather, as follows from the analysis of the covariance matrix (*Eq. 13*), the principal components correspond to the vectors of maximal sensitivity for stimuli of fixed duration.

Concluding remarks

In the present work we focus on the implications of the observed spatiotemporal RFs for coding and stimulus discrimination. Another interesting set of questions involves the origin of the spatiotemporal structure of the response itself. A simple explanation of the structure in terms of the feed-forward neural connections within and beyond the retina is likely to be incomplete. In particular, the dispersion implied by the nonseparability of space and time cannot be readily explained by the properties of individual neurons. A more plausible explanation involves the dynamical response of an interacting network of neurons, possibly amacrine and retinal ganglion cells, whose spatial RFs overlap.

APPENDIX A: SINGULAR VALUE DECOMPOSITION

We consider the expansion of the RF in terms of its SVD (*Eq. 4*). The coefficients λ_n and the functions $F_n(\vec{r})$ and $G_n(t)$ are shown to be the eigenvalues and eigenvectors, respectively, of the correlation matrices of the measured response.

The correlation matrix for the spatial modes is

$$\hat{C}(\vec{r}', \vec{r}) = \frac{1}{T}\int_0^T dt R(\vec{r}', t)R(\vec{r}, t) \qquad (A1)$$

Expansion of the $R(\vec{r}, t)$ terms in *Eq. A1* by *Eq. 4* and use of the orthogonality of the temporal modes (*Eq. 5*) gives

[7] This is a consequence of the continuous time dependence and time-translational invariance of $R(t|\vec{r}, t')$.

$$\tilde{C}(\vec{r}', \vec{r}) = \sum_n \lambda_n^2 F_n(\vec{r}) F_n(\vec{r}') \quad (A2)$$

Multiplication of both sides of *Eq.* A2 and use of the orthogonality of the spatial modes (*Eq.* 5) leads to the eigenvalue equation

$$\int d^2r' F_n(\vec{r}') \tilde{C}(\vec{r}', \vec{r}) = \lambda_n^2 F_n(\vec{r}) \quad (A3)$$

Note that $\tilde{C}(\vec{r}', \vec{r})$ is a symmetric matrix whose rank is bounded by the number of pixels, L_m^2.

The correlation matrix for the temporal modes is

$$\tilde{C}(t', t) = \int d^2r R(\vec{r}, t') R(\vec{r}, t) \quad (A4)$$

Proceeding as above, the temporal modes satisfy the eigenvalue equation

$$\frac{1}{T}\int_0^T dt' G_n(t') \tilde{C}(t', t) = \lambda_n^2 G_n(t) \quad (A5)$$

where $\tilde{C}(\vec{r}', \vec{r})$ is a symmetric matrix whose rank is bounded by the number of frames, N_T. The rank of both correlations matrices must be equal and thus is bounded by the smaller of L_m^2 or N_T, which is $N_T = 16$ in the present case.

The measured RFs $R(\vec{r}, t)$ contain noise, and thus the correlation matrices have a random component that contributes to their eigenvalue spectrum. The number of significant modes in the decomposition of a given RF could be estimated for fields measured with 16×16–pixel stimuli, for which the response of the unit is confined to a subset of the pixels. We compare the spectrum for a 8×8–pixel region over which the unit responded with a 8×8–pixel region for which there is no apparent response. The later region determines the amplitude of the noisy contribution to the eigenvalue spectrum. The number of significant modes in the decomposition is given by the number of terms in the eigenvalue spectrum whose amplitude is significantly above the noise contribution.

APPENDIX B: REALIZATION OF SPIKE TRAINS

Here we describe our realization of neuronal spike trains under the assumption that the spike statistics of each unit are Poisson, with an inhomogeneous rate given by $Z_\alpha(t)$. For Poisson statistics, the probability density of obtaining a spike train $\Lambda_\alpha(t)$, with k spikes at times $t_1 \cdots t_k$, is

$$P[\Lambda_\alpha(t) | S_\alpha] = P[t_1 \cdots t_k | S(\vec{r}, t)] \quad (B1)$$

$$= \frac{1}{k!}\left[\prod_{i=1}^{k} Z_\alpha(t_i)\right] \exp\left[-\int_0^T dt' Z(t')\right] \quad (B2)$$

Each realization of a set of spike times, $t_1 \cdots t_k$, defines a train. We start by considering the probability, $p_\alpha(t)dt$, that the first spike occurs between the times t and $t + dt$, starting at *time 0*. This probability is equal to the probability that no spikes occur between 0 and t and that a single spike occurs between t and $t + dt$. *Equation* B2 yields

$$p_\alpha(t)dt = \exp\left[-\int_0^t dt' Z_\alpha(t')\right] Z_\alpha(t) \exp\left[-\int_t^{t+dt} dt' Z_\alpha(t')\right] dt$$

$$= \frac{\partial}{\partial t} P_\alpha(t) dt \quad (B3)$$

where

$$P_\alpha(t) \equiv 1 - \exp\left[-\int_0^t dt' Z_\alpha(t')\right] \quad (B4)$$

is the probability generating function for $p_\alpha(t)$.

We construct a spike train with the use of the Monte Carlo method. The time between successive spikes is picked up at random according to the distribution $p_\alpha(t)$. The key to this method is to note that the value of the generating function $P_\alpha(t)$ is monotonic between 0 and 1, and thus the inverse function P_α^{-1} exists. If we pick up a random number RAN from a uniform distribution, the random variable

$$t = P_\alpha^{-1}(\mathrm{RAN}) \quad (B5)$$

will be distributed with probability $p_\alpha(t)$. Recurrent application of *Eq.* B5 leads to a set of spike times, $t_1, t_2, \ldots, t_i, \ldots, t_k$, with $0 \le t_1, \ldots, t_k \le T$. In terms of these times the lth realization of the spike train for the αth stimulus is

$$\Lambda_{\alpha,l}(t) = \sum_{i=1}^{k} \delta(t - t_i) \quad (B6)$$

Recall that the times t_i depend on the stimulus through $Z_\alpha(t)$ (*Eqs.* B3 and B4).

APPENDIX C: MUTUAL INFORMATION

The reduction in the uncertainty of knowledge of the stimulus given the response that encodes the stimulus is measured by the mutual information between the spike trains and the stimuli, denoted $I(\Lambda; S)$ (APPENDIX E). It is bounded by $I(\Lambda; S) \le \log_2 N_W$ or $I(\Lambda; S) \le 7$ bits for the set of Walsh patterns. Technically, the mutual information (Cover and Thomas 1991) between the spike trains and the stimuli, $I(\Lambda; S)$, is the difference between the entropy of the train, $H(\Lambda)$, and the conditional entropy of the train given knowledge of the stimuli, $H(\Lambda | S)$, i.e.

$$I(\Lambda; S) \equiv H(\Lambda) - H(\Lambda | S) \quad (C1)$$

In terms of experimental quantities, this becomes

$$I(\Lambda; S) = \sum_i p(\Lambda_i) \log_2 p(\Lambda_i) - \sum_{\alpha=1}^{N_W} p(S_\alpha) \sum_i p(\Lambda_i | S_\alpha) \log_2 p(\Lambda_i | S_\alpha)$$

$$= \sum_i p(\Lambda_i) \log_2 p(\Lambda_i) - \frac{1}{N_W} \sum_{\alpha=1}^{N_W} \sum_i p(\Lambda_i | S_\alpha) \log_2 p(\Lambda_i | S_\alpha) \quad (C2)$$

where $\Lambda_i(t)$ is a particular spike train, $p(\Lambda_i)$ is the probability distribution of the spike trains, $p(\Lambda_i | S_\alpha)$ is the conditional probability of $\Lambda_i(t)$ given knowledge of the stimulus $S_\alpha(\vec{r}, t)$, and the index i extends over all spike trains (APPENDIX D). Further

$$p(\Lambda_i) = \sum_{\alpha=1}^{N_W} p(S_\alpha) p(\Lambda_i | S_\alpha) = \frac{1}{N_W} \sum_{\alpha=1}^{N_W} p(\Lambda_i | S_\alpha) \quad (C3)$$

The space of spike trains is of infinite dimension. We approximate the mutual information over this space by $I(\Lambda; S) \simeq I(\Lambda_1, \Lambda_2, \Lambda_3; S)$, where Λ_n is the projection of the spike train into the subspace spanned by the nth principal component, i.e.

$$\Lambda_n \equiv \frac{1}{T}\int_0^T dt \Lambda(t) \Phi_n(t) \quad (C4)$$

Equation C2 shows that the mutual information $I(\Lambda_1, \Lambda_2, \Lambda_3; S)$ is calculated from the conditional probability $p(\Lambda_1, \Lambda_2, \Lambda_3 | S)$. This conditional probability is calculated with the use of the Monte-Carlo method. For each stimulus S_α, 10^5 spike trains are produced as described in APPENDIX B. The first three principal components $\Lambda_1, \Lambda_2, \Lambda_3$ are computed for each realization with the use of *Eq.* C4. The 3-dimensional space of $\Lambda_1, \Lambda_2, \Lambda_3$ is divided into $22^3 \simeq 10^4$ boxes, and a histogram of the number of realizations falling inside each box for a specific stimulus is calculated. The probability that the response falls inside the bin is the number of realizations in which the response is inside the bin divided by the number of total realizations of the particular stimulus. Because the number of realization is finite, the mutual information calculated with the use of the Monte-Carlo method tends to bias

upward (Optican et al. 1991). Thus a large number of realization is needed. It is shown (Carlton 1969) that the bias in the mutual information calculated this way is proportional to the number of bins divided by the number of trials. Hence the number of trials should be much larger than the number of boxes. We verified that the result is not dependent on the number of boxes by repeating the Monte-Carlo simulation with an eight-times larger number of boxes.

The mutual information between the stimulus and the first principal component, $I(\Lambda_1|S)$ (*Eq.* C4), is calculated by a similar method as is the mutual information between the stimulus and the total number of spikes, $I(\bar{\Lambda}_1 S)$, where $\bar{\Lambda}$ is found by integrating the spike train, i.e.

$$\bar{\Lambda} = \frac{1}{T}\int_0^T dt\Lambda(t) \tag{C5}$$

We now address the dependence of the mutual information on a change in parameters of the system. The neuronal output depends on the background firing rate of the neuron, Z_0, as well as details of the stimuli, such as the contrast modulation. We estimate the effect of changing these factors by parameterizing the average neuron response (*Eq.* 2) as

$$Z_\alpha(t) = g\left[aZ_0 + b\int d^2r\int_{-\infty}^{t} dt'R(\vec{r}, t-t')S_\alpha(\vec{r}, t')\right] \tag{C6}$$

where $a = 1$ and $b = 1$ under normal conditions.

We consider first variations in the background rate and hold b constant. When the background rate in relatively small, i.e., $a \ll 1$, but $b = 1$, the unit operates close to threshold, and we observe that $I(\Lambda; S)$ increases with increasing a. At a critical value of a, typically just below 1, $I(\Lambda; S)$ reaches maximum and then decreases with increasing a. The initial increase occurs because the threshold effect is strong and many stimuli lead to a suppression of activity. The later decrease occurs because a high background rate leads to a higher variance for the neuronal response (*Eq.* 20). The decrease scales as $a^{-1/2}$ in the linear limit but is weaker in practice because of nonlinear effects. Similar results and arguments hold for $I(\bar{\Lambda}; S)$.

Changes in contrast affect the stimulus-related modulation and are modeled by varying the parameter b with $a = 1$. At low levels of contrast, i.e., $b \ll 1$, $I(\Lambda; S)$ increases linearly with increasing contrast. At intermediate levels, but typically with $b < 1$, the probability grows only slowly and with diminishing slope. The details of the growth vary between units. Similar behavior is observed for $I(\bar{\Lambda}; S)$, although the rate of increase is less for a given unit.

Last, we examine the effect of changing the overall gain of the neuron, for which we take $a = b$. An increase in the gain is equivalent to an increase in the number of identical, statistically independent units under the assumption of inhomogeneous Poisson statistics. We observe that both $I(\Lambda; S)$ and $I(\bar{\Lambda}; S)$ increase monotonically with increasing values of a.

APPENDIX D: RESPONSE STATISTICS OF INHOMOGENEOUS POISSON PROCESSES

Assume that a stimulus $S_\alpha(\vec{r}, t)$ creates the response $\Lambda(t)$, so that its trial average is

$$Z(t) = Z_0 + \sum_m \lambda_m A_m G_m(t) \tag{D1}$$

and, for a Poisson process, the variance is

$$\langle\Lambda(t)\Lambda(t')\rangle - \langle\Lambda(t)\rangle\langle\Lambda(t')\rangle = Z(t)\delta(t-t') \tag{D2}$$

where $\langle\cdots\rangle$ signifies trial averaging. The normalization of the eigenvectors from the singular value decomposition, $G_m(t)$, is

$$\frac{1}{T}\int_0^T dtG_m(t)G_n(t) = \delta_{mn} \tag{D3}$$

We estimate A_m from a measurement of the response $\Lambda(t)$. The estimator $\hat{A}_m$ is (*Eq.* 19)

$$\hat{A}_m = \frac{1}{\lambda_m T}\int_0^T dt\Lambda(t)G_m(t) - \frac{Z_0}{\lambda_m}\bar{G}_m = \frac{1}{\lambda_m T}\sum_{i=1}^{k} G(t_i) - \frac{Z_0}{\lambda_m}\bar{G}_m \tag{D4}$$

where the baseline level Z_0 is known and $\bar{G}_m \equiv \frac{1}{T}\int_0^T dtG_m(t)$. The average of the estimator $\hat{A}_m$ is obtained by substituting the probability density $P[\Lambda(t)|S]$ of receiving k spikes at times $t_1, t_2\cdots t_k$ under the inhomogeneous Poisson assumption (*Eq.* B2). We find

$$\begin{aligned}\langle\hat{A}_m\rangle &= \sum_{k=0}^{\infty}\frac{1}{k!}\int dt_1\cdots dt_k\left[\prod_{i=1}^{k} Z(t_i)\right]e^{-\bar{z}}\left[\frac{1}{\lambda_m T}\sum_{i=1}^{k} G_m(t_i) - \frac{Z_0}{\lambda_m}\bar{G}_m\right] \\ &\quad - -\frac{Z_0}{\lambda_m}\bar{G}_m + \frac{1}{\lambda_m T}\int_0^T dtZ(t)G_m(t)\left[\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\bar{Z}^{k-1}e^{-\bar{z}}\right] \\ &= A_m\end{aligned} \tag{D5}$$

The correlation matrix is

$$\begin{aligned}\langle\hat{A}_m\hat{A}_n\rangle &= \sum_{k=0}^{\infty}\frac{1}{k!}\int dt_1\cdots dt_k\left[\prod_{i=1}^{k} Z(t_i)\right]e^{-\bar{z}} \\ &\quad\times\left[\frac{1}{\lambda_m T}\sum_{i=1}^{k} G_m(t_i) - \frac{Z_0}{\lambda_m}\bar{G}_m\right]\left[\frac{1}{\lambda_n T}\sum_{j=1}^{k} G_n(t_i) - \frac{Z_0}{\lambda_n}\bar{G}_n\right] \\ &= \frac{1}{\lambda_m\lambda_n T^2}\int_0^T dtZ(t)G_m(t)G_n(t) + A_mA_n\end{aligned} \tag{D6}$$

The covariance matrix (*Eq.* 20) is

$$\begin{aligned}\sigma^2_{mn}(\{A_m\}) &= \langle\hat{A}_m\hat{A}_n\rangle - \langle\hat{A}_m\rangle\langle\hat{A}_n\rangle \\ &= \frac{1}{\lambda_m\lambda_n T^2}\int_0^T dtZ(t)G_m(t)G_n(t)\end{aligned} \tag{D7}$$

For the case of weak modulation, $Z(t) \approx Z_0$ and the covariance matrix becomes (*Eq.* D3)

$$\sigma^2_{mn}(\{A_m\}) \simeq \frac{Z_0}{\lambda_m\lambda_n T^2}\int_0^T dtG_m(t)G_n(t) = \frac{Z_0}{\lambda_m^2 T}\delta_{mn} \tag{D8}$$

We thank H. S. Seung and H. Sompolinsky for discussions critical to this work and W. S. Bialek, R. Desimone, J. A. Hertz, L. M. Optican, Y. Prut, B. J. Richmond, R. A. Stepnoski, and M. Stryker for useful comments. D. Kleinfeld and B. I. Shraiman thank the Institute of Theoretical Physics, University of California at Santa Barbara, for hospitality.

D. Kleinfeld acknowledges support from the US-Israel Binational Science Foundation, Grant 90-003211/3. R. C. Reid acknowledges support from the National Eye Institute, Grant EY-10115, and the Klingenstein Foundation. R. M. Shapley acknowledges support from the National Eye Institute, Grant EY-01472.

Address for reprint requests: D. Kleinfeld, AT&T Bell Laboratories, 600 Mountain Ave., 1C-463, Murray Hill, NJ 07974.

Received 6 January 1994; accepted in final form 19 July 1994.

REFERENCES

ATICK, J. J. Could information theory provide an ecological theory of sensory processing? *Network* 3: 213–251, 1992.

BIALEK, W., RIEKE, F., DE RUYTER VAN STEVENINCK, R. R., AND WARLAND, D. Reading a neural code. *Science Wash. DC* 252: 1854–1857, 1991.

DE BOER, E. AND KOYPER, P. Triggered correlation. *IEEE Trans. Biomed. Eng.* VBE-15: 169–179, 1968.

BRITTEN, K. H., SHADLEN, M. N., NEWSOME, W. T., AND MOVSHON, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12: 4745–4765, 1992.

CARLTON, A. G. On the bias of information estimates. *Psychol. Bull.* 71: 108–109, 1969.

CHEE-ORTS, M.-N. AND OPTICAN, L. M. Cluster method for analysis of transmitted information in multivariate neuronal data. *Biol. Cybern.* 69: 29–35, 1993.

COVER, M. C. AND THOMAS, J. A. *Elements of Information Theory.* New York: Wiley, 1991.

DAWIS, S., SHAPLEY, R. M., AND TRANCHINA, D. Receptive field organization of X-cell in the cat: spatiotemporal coupling and asymmetry. *Vision Res.* 24: 549–561, 1984.

ENROTH-CUGELL, C. AND ROBSON, J. G. The contrast sensitivity of retinal ganglion cells of the cat. *J. Physiol. Lond.* 187: 517–552, 1966.

GAWNE, T. J., MCCLURKIN, J. W., RICHMOND, B. J., AND OPTICAN, L. M. Lateral geniculate neurons in behaving primates. III. Response predictions of a channel model with multiple spatial-to-temporal filters. *J. Neurophysiol.* 66: 809–823, 1991.

GEISLER, W. S., ALBRECHT, D. G., SALVI, R. J., AND SAUNDERS, S. S. Discrimination performance of single neurons: rate and temporal-pattern information. *J. Neurophysiol.* 66: 334–362, 1991.

GOLUB, G. H. AND KAHAN, W. Calculating the singular values and pseudo-inverse of a matrix. *SIAM Numerical Analysis* 2: 202–224, 1965.

HERTZ, J. A., KJÆR, T. W., ESKANDAR, E. N., AND RICHMOND, B. J. Measuring natural neural processing with artificial neural networks. *Int. J. Neural Syst.* 3, *Suppl.* 1992: 91–103.

HOCHSTEIN, S. AND SHAPLEY, R. M. Linear and nonlinear spatial subunits in Y cat retinal ganglion cells. *J. Physiol. Lond.* 262: 265–284, 1976.

JAGADEESH, B., WHEAT, H. S., AND FERSTER, D. Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science Wash. DC* 262: 1901–1904, 1993.

JONES, J. P. AND PALMER, L. A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* 58: 1233–1258, 1987.

MCCLURKIN, J. W., GAWNE, T. J., OPTICAN, L. M., AND RICHMOND, B. J. Lateral geniculate neurons in behaving primates. II. Encoding of visual information in the temporal shape of the response. *J. Neurophysiol.* 66: 794–808, 1991a.

MCCLURKIN, J. W., GAWNE, T. J., RICHMOND, B. J., OPTICAN, L. M., AND ROBINSON, D. L. Lateral geniculate neurons in behaving primates. I. Responses to two-dimensional stimuli. *J. Neurophysiol.* 66: 777–793, 1991b.

MCCLURKIN, J. W., OPTICAN, L. M., RICHMOND, B. J., AND GAWNE, T. J. Concurrent processing and complexity of temporally encoded neuronal messages in visual perception. *Science Wash. DC* 253: 675–677, 1991c.

MCCLURKIN, J. W., ZARBOCK, J. A., AND OPTICAN, L. M. Temporal codes in monkey striate cortex for colors, patterns and memories. In: *Primary Visual Cortex of Primates. Cerebral Cortex,* edited by A. Peters and K. S. Rockland. New York: Plenum, 1994, vol. 10, p. 443–467.

MCLEAN, J. AND PALMER, L. A. Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Res.* 29: 675–679, 1989.

MILLER, E. K., LI, L., AND DESIMONE, R. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J. Neurosci.* 13: 1460–1478, 1993.

MOVSHON, J. A., THOMPSON, I. D., AND TOLHURST, D. J. Spatial summation in the receptive field of simple cells in the cat's striate cortex. *J. Physiol. Lond.* 283: 53–77, 1978.

OPTICAN, L. M., GAWNE, T. J., RICHMOND, B. J., AND JOSEPH, P. J. Unbiased measures of transmitted information and channel capacity from multivariate neuronal data. *Biol. Cybern.* 65: 305–310, 1991.

OPTICAN, L. M. AND RICHMOND, B. J. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J. Neurophysiol.* 57: 162–178, 1987.

PALMER, L. A., JONES, J. P., AND STEPNOSKI, R. A. Striate receptive fields as linear filters: characterization in two-dimensions of space. In: *Vision and Visual Dysfunction. The Neural Basis of Visual Behavior,* edited by A. G. Leventhal. New York: Macmillan, 1991, vol. IV.

PINTER, R. B. AND NABET, B. *Nonlinear Vision: Determination of Neural Receptive Fields, Function and Networks.* Boca Raton: CRC, 1992.

PODVIGIN, N. F., COOPERMAN, A. M., AND TCHUEVA, I. V. The space-time properties of excitation and inhibition and wave processes in cat's corpus geniculatum lateralis (in Russian). *Biofizika* 19: 341–346, 1974.

REID, R. C. AND SHAPLEY, R. M. Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature Lond.* 356: 716–717, 1992.

REID, R. C., SOODAK, R., AND SHAPLEY, R. M. Linear mechanisms of directional selectivity in simple cell of cat striate cortex. *Proc. Natl. Acad. Sci. USA* 84: 8740–8744, 1987.

REID, R. C., SOODAK, R., AND SHAPLEY, R. M. Directional selectivity and the spatiotemporal structure of the receptive field of simple cells in cat striate cortex. *J. Neurophysiol.* 66: 505–529, 1991.

RICHMOND, B. J. AND OPTICAN, L. M. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. *J. Neurophysiol.* 57: 147–161, 1987.

RICHMOND, B. J., AND OPTICAN, L. M. Temporal encoding of two-dimensional patterns by single units in primate visual cortex. II. Information transmission. *J. Neurophysiol.* 64: 370–380, 1990.

RICHMOND, B. J., OPTICAN, L. M., PODELL, M., AND SPITZER, H. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response properties. *J. Neurophysiol.* 57: 132–146, 1987.

RICHMOND, B. J., OPTICAN, L. M., AND SPITZER, H. Temporal encoding of two-dimensional patterns by single units in primate visual cortex. I. Stimulus-response relations. *J. Neurophysiol.* 64: 351–369, 1990.

SHAPLEY, R. M., REID, R. C., AND SOODAK, R. Spatiotemporal receptive fields and direction selectivity. In: *Computational Models of Visual Processing,* edited by M. S. Landy and J. A. Movshon. Cambridge, MA: MIT Press, 1991, p. 109–118.

SHRAIMAN, B., GOLOMB, D., REID, R. C., SHAPLEY, R. M., AND KLEINFELD, D. On temporal codes and the spatiotemporal response of neurons in the lateral geniculate nucleus. *Soc. Neurosci. Abstr.* 19: 15, 1993.

SO, S. Y. AND SHAPLEY, R. M. Spatial properties of X and Y cells in the lateral geniculate nucleus of the cat and conduction velocities of their inputs. *Exp. Brain Res.* 36: 533–550, 1979.

SUTTER, E. A practical non-stochastic approach to nonlinear time-domain analysis. In: *Advanced Methods of Physiological Systems Modelling,* edited by V. Z. Marmarelis. Plenum, New York: Plenum, 1987, vol. 1, p. 303–315.

TOVÉE, M. J., ROLLS, E. T., TREVES, A., AND BELLIS, R. P. Information encoding and the responses of single neurons in the primate temporal visual cortex. *J. Neurophysiol.* 70: 640–654, 1993.

VITERBI, A. J. AND OMURA, J. K. *Principles of Digital Communication and Coding.* New York: McGraw-Hill, 1979.

WÖRGÖTTER, F. AND KOCH, C. A detailed model of the primary visual pathway in the cat: comparison of afferent excitatory and intracortical inhibitory connection schemes for orientation selectivity. *J. Neurosci.* 11: 1959–1979, 1991.