JOURNAL OF NEUROPHYSIOLOGY
Vol. 70, No. 2, August 1993. *Printed in U.S.A.*

Information Encoding and the Responses of Single Neurons in the Primate Temporal Visual Cortex

M. J. TOVÉE, E. T. ROLLS, A. TREVES, AND R. P. BELLIS
Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom

SUMMARY AND CONCLUSIONS

1. The possibility of temporal encoding in the spike trains of single neurons recorded in the temporal lobe visual cortical areas of rhesus macaques was analyzed with the use of principal component and information theory analyses of smoothed spike trains. The neurons analyzed had responses selective for faces.

2. Provided that a correction was applied to earlier methods of principal component analysis used for neuronal spike trains, it was shown that the first principal component provides by a great extent the most information, with the second and third adding only small proportions (on average 18.8 and 8.4%, respectively).

3. It was shown that the magnitude of the second and higher principal components is even smaller if the spike train analysis is started after the onset of the neuronal response, instead of before the neuronal response has started. This suggests that variations in response latency are at least a part of what is reflected by the second and higher principal components.

4. The first principal component was correlated with the mean firing rate of the neurons. The second and higher principal components reflected at least partly the onset properties of the neuronal responses, such as response latency differences between the stimuli.

5. A considerable proportion of the information available from principal components 1–3 is available in the firing rate of the neuron.

6. Periods of the firing rate of as little as 50 or even 20 ms are sufficient to give a reasonable estimate of the firing rate of the neuron.

7. Information theory analysis showed that in short epochs (e.g., 50 ms) the information available from the firing rate can be as high, on average, as 84.4% of that available from the firing rate calculated over 400 ms, and 52.0% of that available from principal components 1–3 in the 400-ms period. It was also found that 44.0% of the information calculated from the first three principal components is available in the firing rates calculated over epochs as short as 20 ms.

8. More information was available near the start of the neuronal response, and the information available from short epochs became less later in the neuronal response.

9. Taken together, these analyses provide evidence that a short period of firing taken close to the start of the neuronal response provides a reasonable proportion of the total information that would be available if a long period of neuronal firing (e.g., 400 ms) were utilized to extract it, even if temporal encoding were used. The implications of these and related findings are that, at least for rapid object recognition, each cortical stage provides information to the next in a short period of 20–50 ms, does not utilize temporal encoding, and completes sufficient computation to provide an output to the next stage in this same 20- to 50-ms period.

INTRODUCTION

The inferior temporal cortex of primates contains neurons that appear to be involved in the solution of many perceptual problems, including the perceptual invariances (see Rolls 1992). In the majority of studies in this and other parts of the brain, the measure that has been taken of the response of the neuron is its firing rate. However, it is possible that information about the stimulus present is carried not only by the mean firing rate of the neuron, but also by temporal variations in the spike train of the neuron. In a series of interesting papers, Richmond and Optican (Richmond and Optican 1987, 1990; Optican and Richmond 1987) have applied principal component analysis and information theory to the analysis of spike trains produced in single neurons in the primate temporal and striate cortices by different visual stimuli. The general procedure was to smooth the spike train for a single trial with a Gaussian filter (with a σ of 10 ms initially, made adaptive in later papers) (see Richmond and Optican 1990) to produce a spike-density function. The spike-density function was then sampled every 6 ms over 384 ms starting 20 ms after the onset of the visual stimuli, to produce a 64-point time series for each trial. At least five such time series were collected for each stimulus in the stimulus set. From these time series, the principal components were extracted. The principal components are extracted from all individual responses of a neuron to all stimuli. The principal components form a basis set such that the covariance matrix is diagonal, and are ordered so that each component accounts for more variance than any subsequent one. The response of a neuron to a particular stimulus can then be fully described as a weighted sum of these principal components. The proportion of the variance accounted for by each principal component can be taken as a measure of its importance. In addition, to assess how many of the principal components are significant, a bootstrap randomization technique can be applied (as described under METHODS) (Richmond and Optican 1987).

With the use of these techniques, Richmond and Optican showed that several principal components were required to describe the responses of a typical neuron to the different visual stimuli, which consisted of a set of black and white Walsh patterns. The first principal component was typically correlated with the firing rate of the neuron. The second and higher principal components usually had complex time courses, and Optican and Richmond (1987) therefore suggested that information was contained not only in the magnitude of the firing rate response of a neuron, but also in the temporal pattern with which it fired.

Optican and Richmond (1987) went on to analyze the information contained when this temporal aspect of the encoding was taken into account (cf. Eckhorn and Pöpel 1974, 1975) and obtained evidence that ~107% more information about the stimulus was available if the temporal

pattern of firing of the neuron (as expressed by the 1st 3 principal components) to a stimulus over the 384-ms analysis period was taken into account when compared with the information contained only in the firing rate.

Rolls, in studies of the response of single temporal cortical neurons to visual stimuli including faces (see Rolls 1991, 1992), has not observed very clear evidence for different temporal responses of neurons to different visual stimuli. Also, Rolls, and Thorpe and Imbert (1989) have come to the conclusion that, at least for object recognition, information as shown by the response latencies of single neurons flows forward rapidly through the visual system, with response latencies of 80–100 ms being typical in the inferior temporal cortex. (In fact, in the recordings made here, many of the neurons were in the cortex in the superior temporal sulcus a little more posterior than in some of our earlier studies, and the response latencies were in some cases as short as 70 ms.) Given that response latencies under similar conditions are typically 30–50 ms in the striate, primary, cortex, and that there are approximately four cortical stages from striate or V1 to inferior temporal (V1, V2, V4, posterior inferior temporal cortex, anterior inferior temporal cortex), the processing time per cortical stage may be as little as 15–20 ms. Rolls has also observed that the responses of inferior temporal cortical neurons do not typically start with a period of firing that is nonspecific for the different stimuli, and only later become specific. Instead, the responses of the neurons tend by observation to become selective for the different stimuli to which the neuron is tuned very early in the response period.

We therefore performed the systematic investigation described here of the temporal response properties of temporal cortical neurons to visual stimuli. One aim was to investigate, for a well-described population of temporal cortical neurons when responding to their natural effective stimuli, faces, how many principal components were required to account for their responses over long (we took 400 ms) time periods. A difference from the earlier work in our analysis of the information carried by the different principal components is that we made a necessary correction to some of the early principal component analysis methods, as described below and noticed also more recently by Optican et al. (1991). A second aim was to try to identify what might account in part for the second and higher principal components of the temporal aspects of the spike train. A third aim was to measure the information available in different time epochs of the response. In particular, we were interested in the amount of information that could be obtained from a short period (e.g., 20 or 50 ms) of the response of a neuron near the beginning of its response onset, because this is the period within which the information would appear to be required to be present to influence the next cortical stage in time for the delay introduced by each stage to amount to only 20–40 ms. We measured this information, and compared it with the information that theoretically could be extracted if there were 400 ms of analysis time available per cortical stage, which is the order of the period (384 ms) used by Richmond and Optican for their calculations. The information considered in this paper is the information about the stimulus and about the fixation position on the stimulus. The behavioral state was maintained constant for the different stimuli by use of a visual fixation task. This investigation is one of a series (Rolls 1992) designed to investigate the normal functions of the temporal lobe visual cortical areas, and how damage to these brain regions may underlie the perceptual deficits found in patients with damage to these regions.

METHODS

Recording techniques

The activity of single neurons was recorded with glass-insulated tungsten microelectrodes (after Merrill and Ainsworth 1972, but without the platinum plating) in two alert macaque monkeys (*Macaca mulatta;* weight, 3.0 kg) seated in a primate chair with the use of techniques that have been described previously (Rolls et al. 1976, 1990). All procedures, including preparative and subsequent ones, were carried out in accordance with the Policy Regarding the Care and Use of Animals approved by the Society for Neuroscience and were licensed under the UK Animals (Scientific Procedures) Act 1986. The action potentials of single cells were amplified by the use of techniques described previously (Rolls et al. 1979), were converted into digital pulses by the use of the trigger circuit of an oscilloscope, and were analyzed on-line by the use of a MicroVaxII computer. The computer collected peristimulus rastergrams of neuronal activity for each trial and displayed, printed, and stored each trial, as well as computing the peristimulus time histogram by summing trials of a given type. Eye position was measured to an accuracy of 0.5° with the search coil technique, and fixation of the visual stimuli was ensured by use of a blink version of a visual fixation task in which the fixation spot was blinked off 200 ms before the test stimulus appeared. The duration of the test stimuli was 500 ms. After this, the fixation spot reappeared, and when it dimmed after a further random period, the monkey could lick to obtain fruit juice. The stimuli were static visual stimuli subtending 13.5° in the visual field presented on a video monitor at a distance of 1.0 m. One of five fixation spot positions was chosen in random sequence for each trial. The fixation spot positions were at the center of the screen, and in the direction of each corner at a distance of 9.7° from the center of the stimulus. The positions of the fixation points away from the center of the face were located approximately at the edge of the image of the face.

X-radiographs were used to locate the position of the microelectrode on each recording track relative to permanently implanted reference electrodes and bony landmarks. The position of cells is then reconstructed from the X-ray coordinates taken together with serial 50-μm histological sections that showed the reference electrodes and microlesions made at the end of some of the microelectrode tracks (Feigenbaum and Rolls 1991).

Stimulus presentation

Visual stimuli were stored in digital form on a computer disk and displayed on a monochrome video monitor with the use of a video framestore (Advanced Electronic Design 512). The resolution of these images was 256 wide by 256 high with 256 gray levels. The monitor provided maximum and minimum luminances of 6.0 and 0.13 ft-L, respectively, and was adjusted internally and by use of a lookup table for linearity to within 3% with the use of a photometer. The computer randomized the order of presentation of these stimuli, switched the stimuli on and off for each trial, and synchronized its data collection so that the stimulus was turned on at the start of the 21st bin of the peristimulus time histogram.

When digitized visual stimuli were being presented on the video monitor, one set of 4–12 visual stimuli was used at a time. Each set of stimuli was designed to provide neuronal response data relevant

to one or several hypotheses. For example, one set included five different faces, to test whether the neuron responded differently to different faces, and some nonface stimuli such as a sine wave grating, a boundary curvature descriptor, and a complex visual image (see Baylis et al. 1985, Fig. 1), to provide an indication of whether the neuron responded differently to face and to nonface stimuli. The computer randomized the sequence in which the members of the set were presented, and, after it had presented the sequence once, it restarted the set with another random sequence. The computer was allowed to repeat the set 4–10 times to provide sufficient data for an analysis of variance so as to determine whether the neuron responded differently to the different stimuli within the set.

Procedure

As tracks were made into the cortex in the superior temporal sulcus, the responses of each neuron were measured to a standard digitized set of stimuli of different faces and of nonface stimuli (Baylis et al. 1985). If a neuron responded to one or more of the faces, but to none of the nonface stimuli in the set, then a wide range of digitized and real three dimensional (3D) nonface stimuli were shown, to determine whether the response of the neuron was selective for faces. The criterion was that the response to the optimal face stimulus should be more than twice as large as to the optimal nonface stimulus. [In fact, the majority of the neurons in the cortex in the superior temporal sulcus classified as showing responses selective for faces responded much more specifically than this. For half these neurons, their response to the most effective face was more than 5 times as large as to the most effective nonface stimulus, and for 25% of these neurons, the ratio was greater than 10:1. These ratios show that, although responding preferentially to faces, these neurons do not have absolute specificity for faces. Further information on and discussion of the extent to which these neurons have selective responses is given by Baylis et al. (1985). The nonface stimuli from which the optimal was chosen included sine-wave gratings, boundary curvature descriptors, complex 2D stimuli, and complex 3D junk objects, as described above.] If the neuron satisfied the criterion, then a series of 4 face stimuli, which included very effective to noneffective stimuli, was presented in random sequence in the blink task, and then the set was repeated in a new random sequence for a further 4–10 repetitions of the set, to obtain data to allow the responses to each member of the set of stimuli to be analyzed. (With 4 stimuli and fixation in each of 5 positions, there were 20 stimulus conditions, so that there was adequate opportunity for information about the stimulus to be reflected by temporal encoding.) When analysis was performed of information present about both which image was shown and where the fixation position was, there were 4–10 repetitions of each stimulus condition. When analysis of information present about only which image was shown was performed, we were able, given the considerable translation invariance of these cells, to combine together trials with different fixation positions, so that in this case there were 20–50 repetitions of each condition. The face stimuli were monkey or human, usually in frontal view, but in some cases in profile view. The receptive fields of the cells were sufficiently large that for most cells there was only a small alteration in the average firing rate when fixation was at one of the positions that was 9.7° from the center of the screen (E. T. Rolls, M. J. Tovée, and P. Azzopardi, unpublished observations).

Data analysis

The principles of the data analysis were similar to those developed by Richmond and Optican (1987) and Optican and Richmond (1987), except that we applied a correction procedure for the limited number of trials similar to but not identical to that developed by Optican et al. (1991), in a way made clear below. A novel aspect of the data analysis performed here is that we investigated how much information was available in short epochs of the spike train.

REPRESENTATION OF NEURAL RESPONSES. The raw data, expressed as the poststimulus occurrence times of individual spikes, recorded with 1-ms resolution, was smoothed by convolution with a Gaussian kernel with a σ of 5 ms (corresponding to low-pass filtering with −3-dB cutoff at 240 Hz). The smoothed data was then quantized, for some of the later analyses, into bins of width 10 ms. As a result the numbers of spikes in each bin *i*, giving the components x_i of a multidimensional response vector, were real rather than integers.

EXTRACTION OF THE PRINCIPAL COMPONENTS OF THE VARIANCE. To analyze the temporal course of the response, a certain number B of consecutive bins (typically, $B = 40$) were selected to form a basis in the space of response vectors. In most cases we chose the bins to cover either of the two poststimulus time periods 0–400 or 150–550 ms. If we label the bins with the subscript i (e.g., i = 1, 40), the response vector during a particular trial k is denoted as x_i^k. The covariance matrix of the responses recorded from a particular cell is

$$C_{ij} = \sum_k (x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j)/N$$

where N is the total number of trials for that cell, and x_i is the average number of spikes, over those N trials, falling in bin *i*. Eigenvalues and eigenvectors of the covariance matrix were extracted by the use of a standard algorithm. The eigenvectors, labeled i = 1,.., B in order of decreasing eigenvalue, form a new basis for the response space. Each subtracted response vector (i.e., with the average response vector subtracted out) is now expressed in this new basis via multiplication by the appropriate orthogonal matrix. The resulting coefficients c_i^k [sometimes called the Karhunen Loeve transform of the vector $(x_i^k - \bar{x}^k)$] are uncorrelated, in the sense that their covariance matrix is, by construction, diagonal. Note that, although the eigenvectors of C_{ij}, the principal components, are normalized to unit length, to bring out their relative contribution to the variance they are plotted, in the figures (e.g., Fig. 4), after multiplication by the square root of the corresponding eigenvalue.

MEANS AND STATISTICAL SIGNIFICANCE OF THE COEFFICIENTS. The coefficients c_i^k were then averaged over the trials with the same stimulus. (Stimulus refers here to 1 experimental stimulus condition. The number of experimental conditions in these experiments was the number of images × the number of fixation positions, i.e., 20.) The means c_i^s represent the loading of stimulus s on component i, and their statistical significance can be tested with the use of the bootstrap method (Richmond and Optican 1987). A scrambled stimulus-response pairing is obtained by assigning random selections of actually recorded responses to new stimulus labels, while maintaining the same number of trials per stimulus as in the authentic pairing. The coefficients c_i^s are then calculated for each pseudostimulus label, with the number of labels high enough that the distribution of coefficients approach, for each component, a continuous distribution (we used 100 times as many labels for the bootstrap distribution as experimental stimuli). The statistical significance of the original c_i^s can now be tested against the null hypothesis distribution that the relationship between stimulus and response be purely random. We note that if there are 40 coefficients, then the response space is 40-dimensional. However, given that there were 20 stimulus conditions, if there were no noise (i.e., the response were uniquely determined by the stimulus), we would expect no more than 19 components to be significant.

RAW INFORMATION MEASURES. If S denotes the set of all stimuli s, and R the set of responses, the average information contained in the responses of a particular cell about that set of stimuli is defined as

$$I(S, R) = \sum_{s \in S} \sum_{r \in R} P(s, r) \log_2 \frac{P(s, r)}{P(s)P(r)}$$

where P is the probability of occurrence of a particular event. In evaluating the information content from the data recorded, we consider different definitions of the response set. In one case, responses are simply quantified by the number of spikes within a preset time period (a unidimensional measure). In another case we consider a multidimensional response space spanning the first Q (typically, $Q = 1, 2$, or 3) of the principal components extracted with the algorithm above. If, for example, three principal components are included, the particular response relative to trial k is defined as the triplet (c_1^k, c_2^k, c_3^k). Although the set of stimuli can be discrete (it is in the present experiment), R is generally a continuum (in the 2nd case, a Q-dimensional vector) space. Because, in practice, one has to evaluate the expression for I by performing a sum rather than an integral, R needs to be quantized. We perform this second quantization following a procedure similar to the first. Consider first the spike count case. The original data are represented by the number of spikes n^k recorded in trial k within the prescribed window, minus the average over all trials. The range of the data is set as $(-\Delta n, \Delta n)$, with Δn the largest between the maximum and absolute value of the minimum recorded for that cell. This range is divided into a preselected number D of bins (we ultimately used $D = 15$). $D - 2$ bins have width $dn = 2\Delta n/(D - 3)$, with the second bin centered around $-\Delta n$, the third shifted by dn, and so on. The first and last bins cover the two semi-infinite intervals at the extremes. If we consider now only the trials relative to a given stimulus s, a smoothing procedure is applied by convolving the individual values n with a Gaussian kernel of width the standard deviation $\sigma_n(s)$ of the values relative to the same stimulus. The result, normalized by dividing by the total number of trials, is quantized into the bins defined above, the area within each bin being used as an estimate of the joint probability $P(s, r)$, where r corresponds to one of the response bins. Summing over all stimuli gives $P(r) = \sum_{s \in S} P(s, r)$. In the case, instead, of the information contained in the Q principal components, the original data are represented by the Q-dimensional set of coefficients c^k. Note that the average response has already been subtracted out. One repeats the same procedure Q times to find the ranges spanned by the coefficients for each component, divide the ranges into bins, smooth the distribution of values recorded for each stimulus, and find $P(s, r)$. Each of the D^Q bins of the response space is now the product of independently determined bins for each of the Q components, and the relative probability is obtained by multiplying the estimated probabilities of the response falling into each unidimensional interval. As noted by Optican and Richmond (1987), it would be preferable to estimate $P(s, r)$ by using the more time-consuming procedure of recalculating separate principal components for each stimulus, because these are, in general, different from those calculated over the whole stimulus set. However, this simplification in the construction of the estimator $P(s, r)$ is of minor import when compared with the distortions produced by limited sampling (see below).

SUBTRACTED INFORMATION MEASURES. The procedure introduced so far for estimating the probability $P(s, r)$ of a particular response is rather simple. Slightly more sophisticated estimators can be used (see Fukunaga 1972; Optican and Richmond 1987) that attempt to take into account the limited number of trials available, in practice, for each stimulus. Ultimately, however, the fact that the probability distribution for each response is extracted from a small sample remains a problem, which cannot be solved by adjusting the kernel used for smoothing, nor by a clever quantization into bins. This can be seen by using again a bootstrap procedure similar to the one used to test the significance of principal component coefficients (Optican et al. 1991). One generates a scrambled stimulus-response pairing, with the number of pseudostimulus labels now equal to the number of real stimuli, and calculates the information I_0 contained in the responses about their randomly paired pseudostimuli. I_0 should be zero, whereas it turns out, in practice, to be of the same order as the original I. This is because the use of probability distributions based on a limited number of trials per stimulus unavoidably biases upward an information estimate, by producing random fluctuations in the distributions of responses to different stimuli, which are then picked up by the information measure as being related to the stimuli.

The most straightforward remedy is to subtract I_0 from I, to obtain a measure $I_s = I - I_0$ that averages zero, even for small samples, in the absence of a stimulus-response causal relationship. This is one of the simplest of several correction procedures for sampling bias in information estimates that have been reported in the literature (see Optican et al. 1991, and references therein). It should be noted that only a few of these procedures are applicable to data for which the a priori distribution is unknown, and that the relative validity of different methods is strongly dependent on the data distribution itself (Macrae 1971). The latter of these considerations, especially, makes us tend to prefer the simplest and most conservative procedure (i.e., $I_s = I - I_0$), rather than the one which consists in subtracting, instead of the fraction of spurious information, its square, i.e., $I'_s = I[1 - (I_0/I)^2]$, as proposed by Optican et al. (1991). We note that subtraction of the square (which is obviously less than the fraction itself) leaves an estimate I'_s often closer to I than to I_s. Optican et al. (1991) argue in favor of subtracting the square on the basis of computer simulations (and data from 2 cells in the striate cortex), which seem to indicate a faster convergence to the asymptotic "correct" value. We note, however, that those computer simulations, based on different distributions than the unknown ones characterizing real data, could lead to inappropriate conclusions with small sample sizes (i.e., numbers of trials), although of course both correction procedures approach the same asymptotic correct value as the sample size increases. We note that, because the upward bias of I increases with the dimensionality of the response space, the use of some subtracted estimate is particularly important when calculating the information contained in the first Q principal components, with $Q > 1$.

Because, on the basis of the above points, it is not clear which correction procedure should be used, we have specifically investigated which correction procedure may be most appropriate for data obtained from real neurons. The results of this analysis applied to 10 of the cells analyzed in this paper are shown in Fig. 1. (The 10 cells included in Fig. 1 are those for which data from 60 presentations of each image were available.) The calculated raw information measure (□) is clearly an overestimate of the true information contained in the neuronal responses about the stimulus, in that this measure decreases sharply as the number of presentations of each stimulus increases, especially when the number of presentations of each stimulus is in the range 3–30. Use of the correction procedure $I_s = I - I_0$ (which we will term correction procedure 2 for reference), shown as filled squares in Fig. 1, although tending to underestimate the information when there are few (3–10) presentations of each stimulus, is likely to be more accurate with 15–60 presentations per stimulus, in that comparison of correction 2 with the raw measure shows that even with 60 presentations per stimulus with our data, I_0 was not zero, so that even with 60 presentations of each stimulus, some correction should be applied. (As soon as I_0 is 0, there is no longer any need for a correction.) In contrast, the measure $I'_s = I[1 - (I_0/I)^2]$

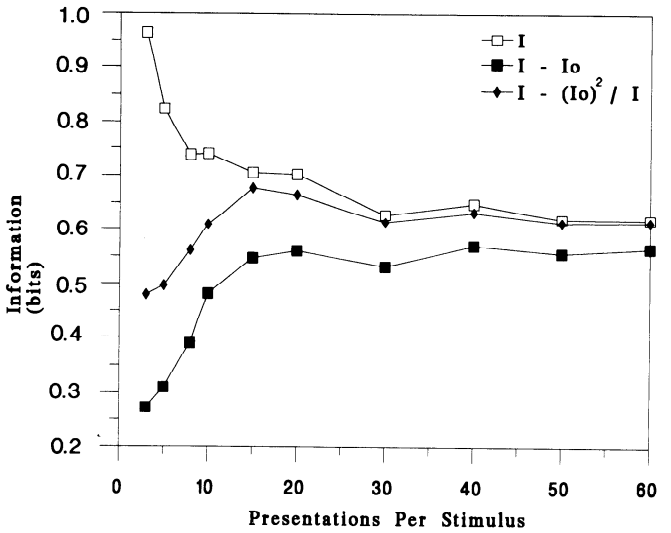

FIG. 1. Graphs showing the information about which stimulus was present (I_s, ordinate), based on the 1st 3 principal components, when no correction was applied (I), when correction procedure 1 $\{I'_s = I[1 - (I_0/I)^2]\}$ was applied, and when correction procedure 2 ($I_s = I - I_0$) was applied, as a function of the number of presentations of each stimulus. I_0 is the spurious information contained in the neuronal responses when the analysis is run with the neuronal responses randomly paired with stimuli.

(which we will refer to as correction 1) underestimated less than correction 2 with small numbers of presentations of each stimulus (3–7) but did provide somewhat of an overestimate of the information about the stimulus for all other numbers of stimulus presentations in the range 10–60. (For example, correction 1 gave information values that approached those given by the raw measure for 30–60 stimulus presentations even though I_0 was not 0, as shown in Fig. 1.) In the light of these findings, we present the data later in Fig. 8 showing the result of applying both corrections, knowing that correction 2 is more conservative. In later figures, the data are shown with correction 1 only, even though this will give a small overestimate of the amount of information available (as shown in Fig. 1), because correction 1 gives a smaller underestimate with small numbers of presentations of each stimulus. Correction 1 also allows comparison with other data (Optican et al. 1991). However, it must be borne in mind later that use of this correction may tend to overemphasize the proportion of information available in principal components higher than 1 (as shown later in Fig. 8). We have presented some of the differences between these two correction procedures and look forward to future investigations with even much larger sample sizes to further clarify which correction is more appropriate.

It should be noted, finally, that an information estimate based on a quantized response set tends to grow with the size D^Q of the set for D small, until it saturates once the width of the D bins along each dimension becomes negligible with respect to the standard deviation of the relative coefficients for each stimulus. We used $D = 15$ (similar to the value of 12 used by Optican and Richmond 1987) after checking that no marked increase in I_s resulted from using larger D values.

RESULTS

It was possible to complete the analyses for 48 neurons recorded in 2 monkeys. The neurons were in the cortex in the superior temporal sulcus or in the inferior temporal visual cortex (see Fig. 2).

Examples of rastergrams and spike-density functions are shown in Fig. 3 for one neuron. Responses to four different stimuli are shown.

The first four principal components extracted from the neuronal responses to 20 different conditions (4 different faces with the fixation spot in 5 different positions for each face) are shown in Fig. 4 (for the neuron shown in Fig. 3). The first principal component accounted for 20% of the variance. It can be seen that this has a similar form to the spike-density function and is thus closely related to the firing rate of the neuron. The second principal component accounted for 8.9% of the variance and had a biphasic component at about the time of the average onset latency of the neuron. This finding suggests that this principal component could be related (at least for this neuron) to the latency of onset of the neuronal response to the different stimuli. This can be understood by realizing that the spike-density functions to different stimuli could be generated by taking the first component with one fixed weight, and the second component with a different weight for each of the different response latencies. In this example, short response latencies would be generated by including the second component with a negative weight (because the 1st part of the principal component has negative sign). Long response latencies to other stimuli could be reconstructed by taking positive components for the second principal component. After the initial biphasic part of the second principal component around 100 ms poststimulus, the remainder of the second principal component is close to zero, indicating that it is not related to the ongoing firing rate of the neuron once it has started to respond to a stimulus. The third and fourth principal components are smaller (6.0 and 4.9%) and could be related to small differences in the time course to different stimuli at different poststimulus times.

To help to interpret the different principal components further, the weighting of the response to a given stimulus and the fixation position on each of the first two principal components is plotted as a function of the overall firing rate of the neuron to that stimulus measured in a 400-ms period starting at stimulus onset (see Fig. 5). Each point in Fig. 5*A* represents the firing rate to one of the stimuli, and the associated weight coefficient of the neuronal response to that stimulus on the first component. It is clear that the stimuli to which the neuron had a high firing rate had a large weighting on the first principal component. The correlation between the rate and the coefficient was 0.99. In contrast, there was no relation between the second (and higher) principal components and the firing rate of the neuron, as shown in Fig. 5*B*. (The correlation in the case of the 2nd principal component was 0.11.) Thus the first principal component (but not higher components) was related to the rate of firing of the neuron, as suggested above and as also found by Richmond and Optican (1987).

Similar analyses were performed for the other cells recorded, and the results for the population analyzed are shown in Fig. 6. Most of the cells had a high correlation between their firing rate and the first principal component (Fig. 6*A*). On the other hand, there were typically only low correlations between the firing rate of the cells and the second principal component (Fig. 6*B*). The same low correlations were typically found for the higher principal components. This analysis shows that the first principal compo-

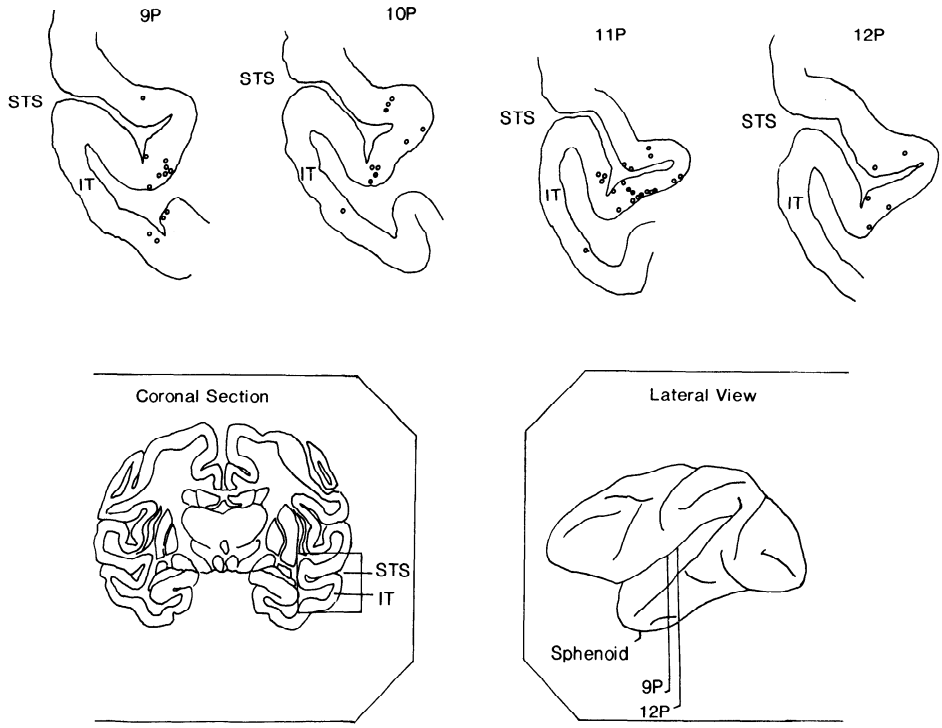

FIG. 2. Recording sites, shown on coronal sections, of the neurons included in this study. Positions of the coronal sections are shown on a lateral view of the macaque brain. The distances refer to millimeter posterior to the sphenoid reference plane (see text).

nent is closely related to, and reflects, the average firing rate of the neuron in the 400-ms analysis period.

To provide some evidence on the information reflected in the second (and perhaps higher) principal components, we analyzed their shape. For a number of neurons, the second principal component was related in time to approximately the onset of firing of the neuron to the stimulus (see, e.g., Fig. 4, *top right*). We therefore analyzed whether there might be systematic onset latency differences for some of the different stimulus conditions. As shown in Fig. 7, latency differences for different stimulus positions on the retina were sometimes present. For example, the neuron shown in Fig. 7 responded with a latency that was ~20 ms longer for fixation at the edge of a face than at the center of the face. The shape of the second principal component for this neuron, and the fact that the second principal component was related to stimulus position more than stimulus type, indicated that it was related to the temporal aspects of the response train only by reflecting the differences in response latencies of the neuron. Indeed, for that neuron the second principal component was similar to the difference of the peristimulus spike-density functions (Fig. 7, *bottom trace*), and, across the set of stimuli, different loadings on the second principal component could have accounted for the response onset latency differences to the different stimuli. For at least a proportion (~50%) of the neurons, the second or a higher principal component could be seen to reflect the onset latency of the neuron in different stimulus conditions, rather than temporal variations in the spike train once the spike train had started. (For the other cells, such differences in response latency may have been present but did not, given the sample size, reach statistical significance.) In addition, it was found that the weightings of the neuronal responses to a given stimulus on the principal components higher than three were not significant (as assessed with the bootstrap procedure) more often than would be expected by chance. Thus, across this population of neurons, it was only the first, and in some cases the second and third, principal components that were consistently related to which stimulus was shown.

We can now consider the information present in the spike train; and in particular how, as principal components above the first are added, the addition adds significantly to the information about the stimulus that can be derived from the spike train (Optican and Richmond 1987). This is a quantitative way of assessing the importance of the higher principal components. The information present in the spike train with the first (1), with the first two (2), and with the first three (3) principal components included is shown in Fig. 8. The separate parts of the figure, *A* and *B*, are for analysis conducted in the 400-ms time period starting either at the onset of the visual stimulus (0–400 ms in Fig. 8*A*), or starting 150 ms after stimulus onset (150–550 ms in Fig. 8*B*). The latter analysis was performed because, with typical neuronal response latencies of 80–100 ms, the onset of the neuronal response is not included, and instead it is possible to assess whether, once the neuron is firing to its effective stimulus, there is still extra information available if the higher, time-varying, principal components are included in the analysis. The amounts of information present when no correction was applied (the raw, uncorrected, information I), when correction procedure 1 $\{I[1 - (I_0/I)^2]\}$ was applied, and when correction procedure 2 ($I_s = I - I_0$) was applied are shown.

It is shown in Fig. 8*A* for the 0- to 400-ms analysis period that, if no correction is applied [the situation considered by Optican and Richmond (1987)], then apparently considerable amounts of information are added by the second and third principal components (for image and position, 62.5% by PC2, and 38.5% by PC3). However, when a correction for limited sampling was applied, most of the information is present in the first principal component, and adding the

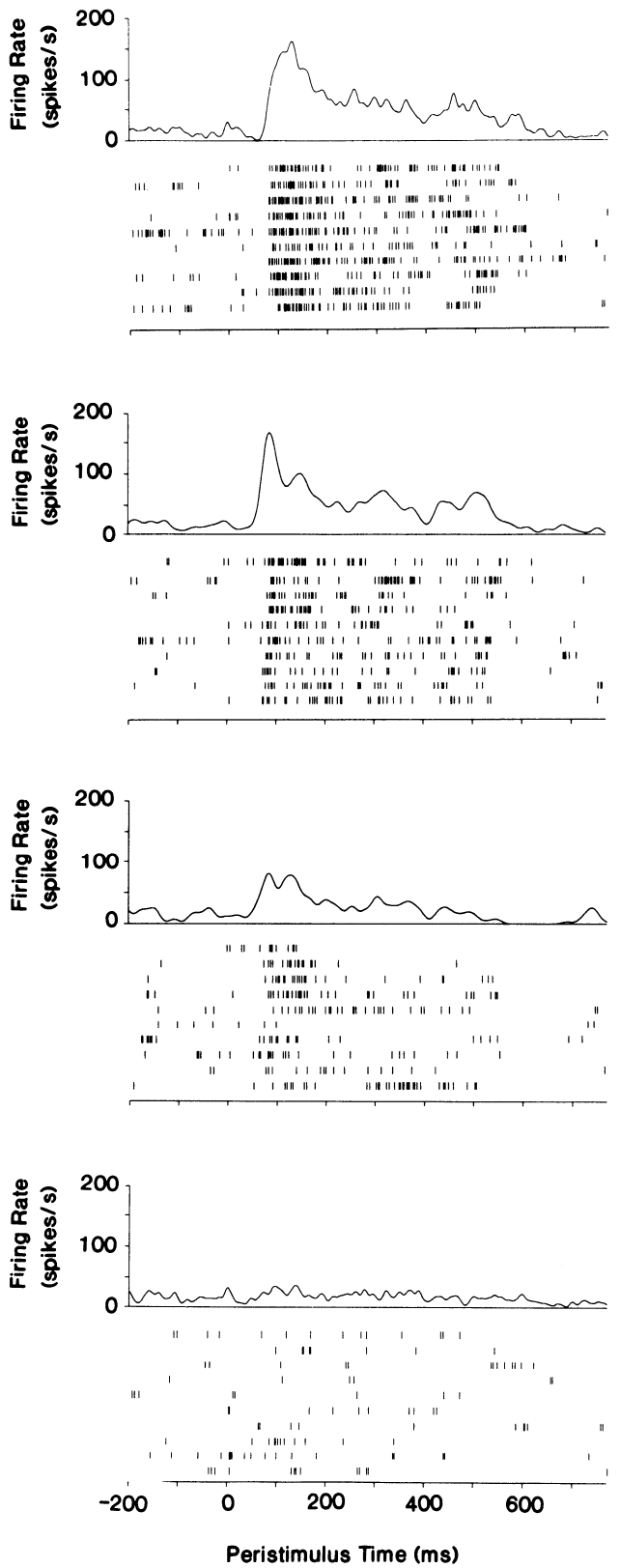

FIG. 3. Peristimulus rastergrams and spike-density histograms showing the response of a single neuron in the inferior temporal cortex to 4 visual stimuli (presented in pseudorandom sequence). The visual stimulus was shown at *time 0*. Each horizontal raster shows the response of the neuron on 1 trial, and each vertical line in a raster represents 1 spike from the cell. The stimulus was visible for 500 ms.

second and third principal components does not add very much more information. (For correction procedure 1, the information added by the 2nd and 3rd principal components was 35.9 and 13.5%, respectively. For the more conservative correction procedure 2, the information added by the 2nd and 3rd principal components was 18.8 and 8.4%, respectively).

It is shown in Fig. 8*B* for the 150- to 550-ms analysis period that, if no correction is applied [the situation considered by Optican and Richmond (1987)], then apparently considerable amounts of information are added by the second and third principal components (for image and position, 60.1% by PC2, and 34.5% by PC3). However, when a correction for limited sampling was applied, most of the information is present in the first principal component, and adding the second and third principal components does not add very much more information. (For correction procedure 1, the information added by the 2nd and 3rd principal components was 26.1 and 13.0%, respectively. For the more conservative correction procedure 2, the information added by the 2nd and 3rd principal components was 13.4 and 8.5%, respectively).

The full height of the histograms in Fig. 8 shows the analysis performed over image and position on the screen and is what has been discussed above. (There were 4 images and 5 positions in which the image could appear on the screen.) The dotted line shows the analysis performed over image only, for comparison. Even more in the case of image alone than in the case of image and position, the conclusion that adding principal components beyond the first adds relatively little further information about the stimulus is the case. (This is probably related to the fact that, as in some cases, the neurons had slightly different response latencies when the stimulus was presented away from the fovea; this onset latency difference could be carried by one of the higher, 2nd or 3rd, principal components.) Because little further information was added by the third and higher principal components, and because the weightings of the responses on principal components higher than three were not consistently significant, we did not normally calculate the information provided by more than the first three principal components, and show the information in the first three principal components when appropriate in the remainder of this article.

The analysis shown in Fig. 8*B* makes it clear that, if the analysis is confined to the period after the cell has started to respond (i.e., the 150- to 550-ms data analysis period shown), then very little extra information is added by the second and third principal components. This is consistent with the view that once a spike train has entered the response period, there is little temporal encoding within it of the nature of the stimulus. However, if the onset of the neuronal response is included in the analysis (Fig. 8*A*), then some extra information is added by the second and higher principal components, because they can then provide information about the response latency of the neuron, which may be different for the different stimulus conditions.

The analysis just described, the results of which are shown in Fig. 8, confirm the results of Optican and Richmond (1987) in showing that information is apparently

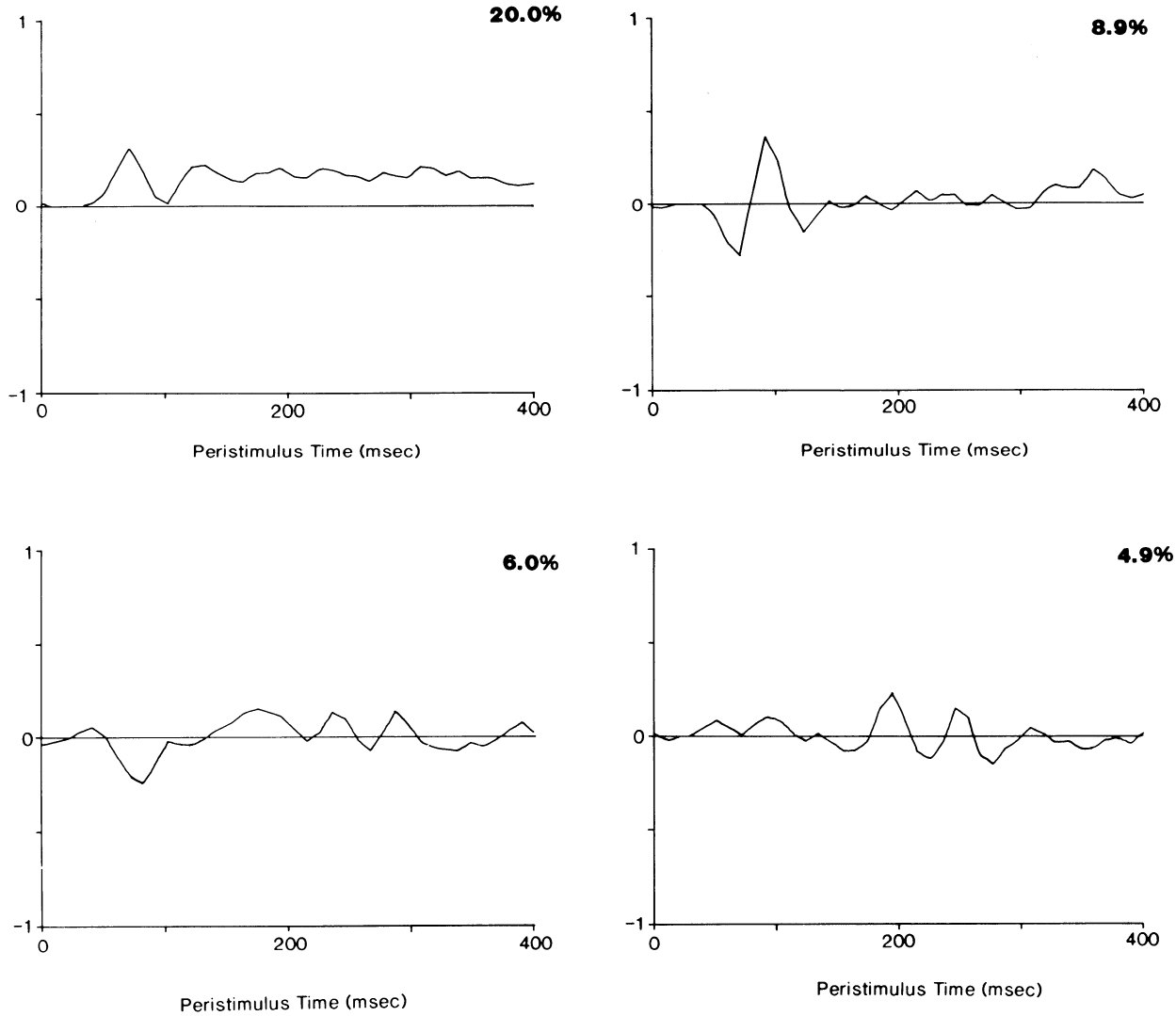

FIG. 4. First 4 principal components extracted from the responses of the neuron shown in Fig. 2 to four different visual stimuli fixated in each of 5 different positions. The percentage of the variance accounted for by each principal component is shown. The scale of the ordinate is constant for the different graphs (see text; Principal Component 1: *top left;* PC2: *top right;* PC3: *bottom left;* PC4: *bottom right*).

added by principal components higher than 1 if the raw, uncorrected, information measure is used. However, the analysis we have just presented shows that much less information is added by the second and higher principal components if the appropriate correction procedures for limited sample sizes are applied, and that it is necessary to apply such correction procedures. The results also suggest that part of what can be reflected in principal components higher than 1 is the onset latency of the neuronal response. Moreover, the results described here confirm that the information contained in the first principal component reflects the firing rate of the neuron. For the rest of the data shown in this paper, we applied correction 1, for the reasons described under methods, and because correction 1 is the same as that applied by Optican et al. (1991), allowing comparisons of the data between studies.

The analysis described here suggests that the major part of the information that can be derived from the spike train is contained in the first principal component, and reflects the firing rate of the neuron in the 400-ms data analysis period. To allow assessment of this point further, the information that can be derived from the firing rate is compared with that which can be obtained from the first three principal components in Fig. 9. The data based on an analysis of the information available about image and position from principal components 1–3 (PC), and from the firing rate are shown, for different periods of analysis of the spike train. The values shown represent the means across all 48 cells. It is shown that, consistent with the points made above, more information about the stimulus is present if PCs1–3 are taken (reflecting temporal encoding) than if firing rate alone is measured, but that much of the information is present in the firing rate. It is also shown that, if the onset of the neuronal response is excluded from the analysis, so that the principal components do not carry information simply because they can include information about neuronal response latency, then the information available from the firing rate becomes closer to that available from the temporal encoding reflected in the PCs1–3.

The data presented so far on the information available in

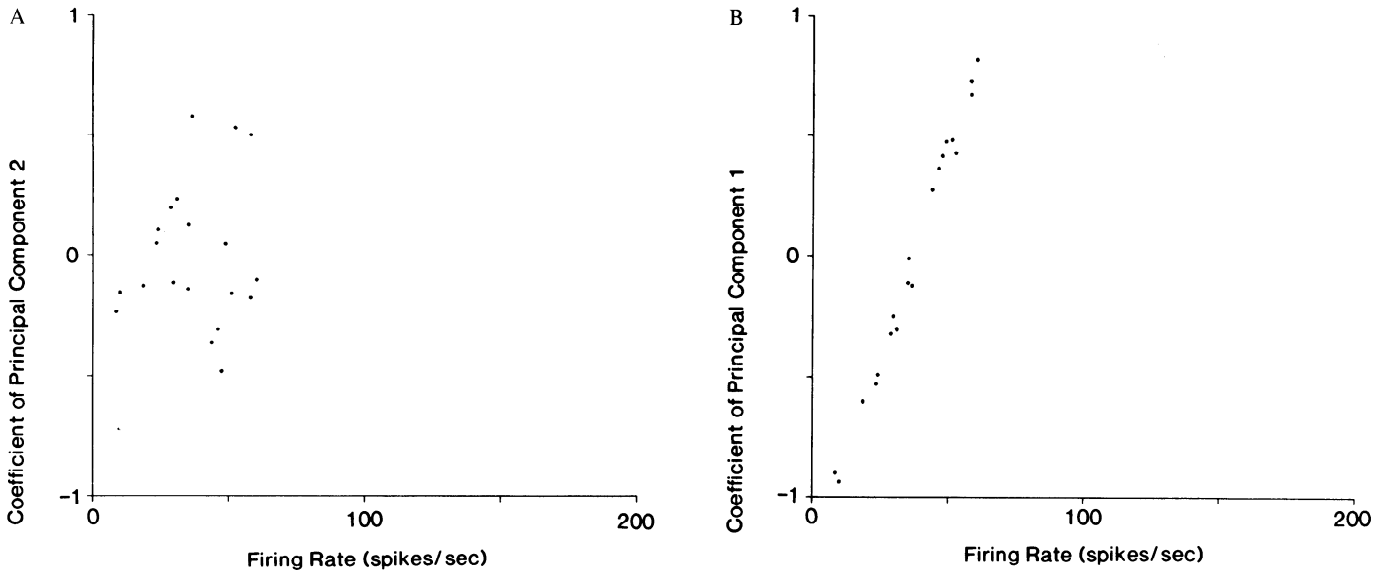

FIG. 5. Each graph shows the relation between the firing rate of the cell shown in Fig. 3 to one of the image/position combinations on the retina and the corresponding weighting (coefficient) on each of the 1st 2 principal components.

the spike train have been analyzed primarily in long data analysis periods, of 400 or 500 ms. We now consider shorter data analysis periods, to provide evidence on the information that would be available from the neuron about the stimulus if a period <400 ms were available for each cortical stage to provide a signal for the next cortical stage.

One way in which we assessed this was by measuring how the mean firing rate of the neuron in a short period was correlated with the rate measured in a long period, e.g., of 500 ms starting 100 ms after stimulus onset. (A period starting with a latency of 100 ms was used because most of the neurons had started to respond within 100 ms of the onset of the stimulus.) It is shown in Fig. 10, *A* and *B*, that in epochs of 100 and 50 ms the correlations with the mean rate in the 500-ms period were high (0.95 and 0.84, respectively). The correlation was lower, although still very significant, when the response was measured over a period of 20 ms (Fig. 10*C*). The regression lines are shown in the figures. The data in these figures are for one cell (the same cell as that shown in Fig. 3). Each point shown is the mean for one of the stimulus conditions (e.g., 1 image in 1 position on the screen).

To assess how adequately these short analysis periods were able to provide a reasonable estimate of the mean firing rate of the cell, the correlations for each of the cells analyzed between the firing rates in the 500- and the 50-ms period are shown in Fig. 11, *right*. It is apparent that for most of the cells the correlation was good. For the majority of the cells, the correlation was also reasonable when only a 20-ms sample of the firing rate was taken (Fig. 11, *left*).

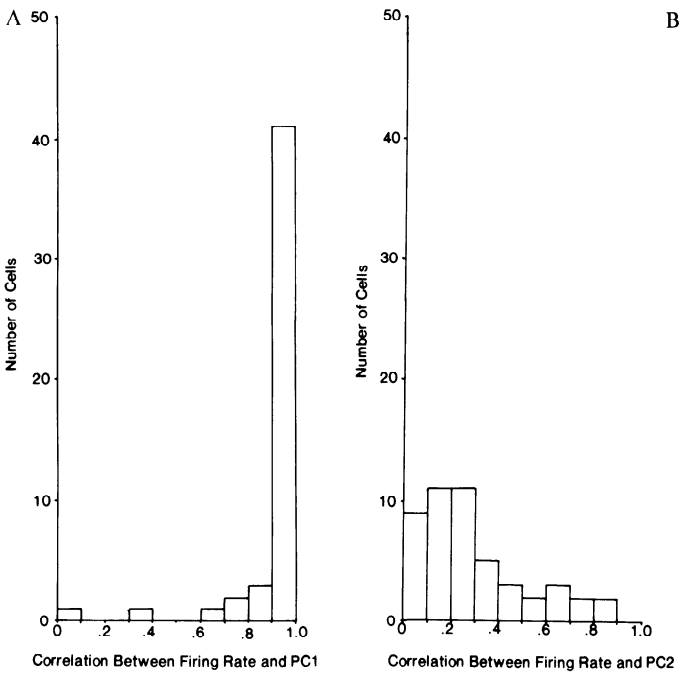

FIG. 6. *A*: histogram showing that for most cells, there was a high correlation between the firing rate of the cell and principal component 1. (The 1st principal component extracted is numbered 1.) *B*: histogram showing that for most cells, there was a low correlation between the firing rate of the cell and principal component 2.

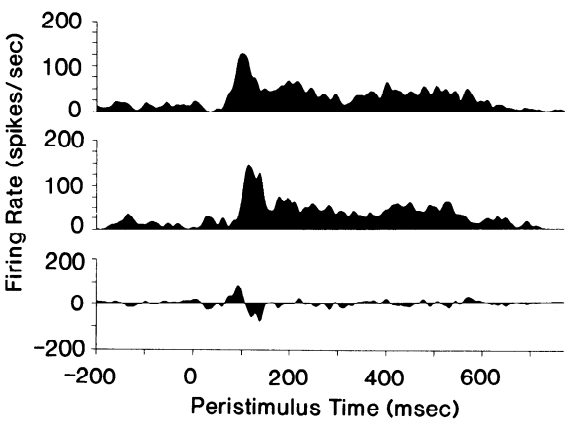

FIG. 7. Spike-density functions for a neuron when the the fixation was at the center of the stimulus (*top spike-density function*) or at the edge of the image (*middle spike-density function*). As shown by the difference between these spike-density functions (*bottom function*), the neuron responded with a latency that was shorter by ~20 ms for fixation at the center as compared with fixation at the edge of the image.

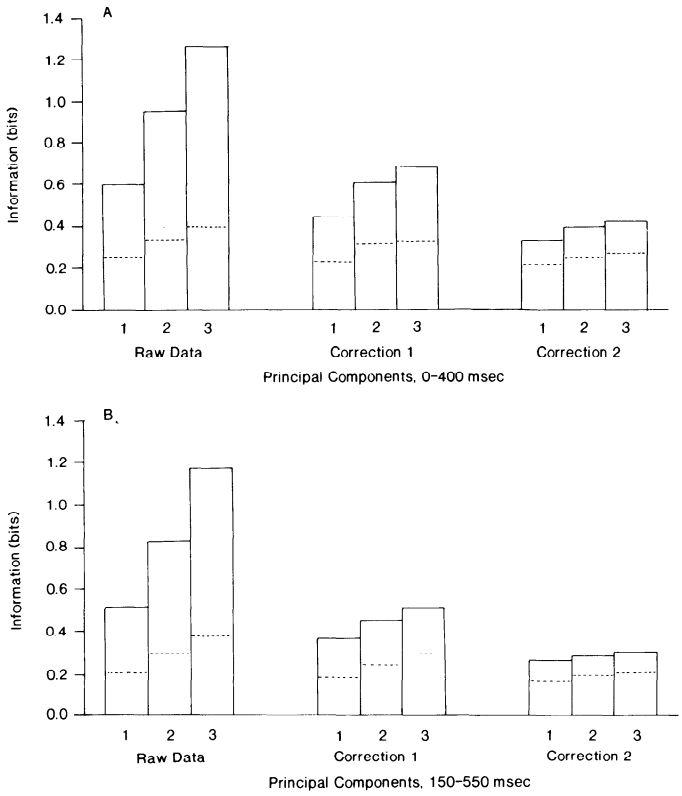

FIG. 8. *A*: for the 0- to 400-ms data analysis period, the information available from the 1st, from the 1st 2, and from the 1st 3 principal components, shown when no correction was applied (*I*, raw data), when correction procedure 1 $\{I'_s = I[1 - (I_0/I)^2]\}$ was applied, and when correction procedure 2 ($I_s = I - I_0$) was applied. Solid bar shows information available about the image and its position; dashed line shows information available about image type only. Data shown are the means over all 48 cells for the information represented by each cell about which image was shown. *B*: same as *A*, except that the data were calculated for the poststimulus period 150–550 ms (which includes the firing once the response has started at ~80 ms, but excludes the onset of the neuronal response).

The data shown in Figs. 10 and 11 indicate that short periods of the firing rate, taken early in the neuronal response period, can give a reasonable estimate of the firing rate taken over a much longer analysis period, of 500 ms. To show quantitatively how much information is present in periods of different durations taken at different times after the stimulus onset, we calculated the information available as described under methods. [It should be noted that for these short response periods, the information must be based on firing rates in the periods, or on principal component analysis utilizing many fewer than the 40 time bins used in the previous analysis. It is useful to understand that the reason for this is that 40 principal components of a time-varying spike train simply cannot be extracted and are essentially meaningless, if the spike train is short, because only 1 or a few time samples would be effectively available as a result of the necessary temporal smoothing required to extract the time series (cf. Richmond and Optican 1987).] The results are shown in Fig. 12. The information based on rates in the 0- to 400-ms period is included to show what is available with a long analysis time. In all cases, the data shown in Fig. 12 are the means over the whole population of 48 cells analyzed and are not just for one cell, so that they give a fair indication of the properties of this population of cells.

First, it is shown in Fig. 12*B* that a 100-ms sample taken starting at 100-ms latency provides almost as much information as is available from the 400-ms rate estimate shown in Fig. 12*A*. (A little more information is available in the 400-ms period starting at 100 ms than in the 400-ms period starting at stimulus onset, because for the 1st part of the 0- to 400-ms period the neuron did not start to respond.) It is also shown in Fig. 12*B* that of the different 100-ms periods, most information was present in the period from 100–200 ms (when the neurons were firing fastest).

Second, it is shown in Fig. 12*A* that 50-ms samples provide a reasonable amount of information, provided that the sample is taken at the right time, which is early in the response period, in particular 100–150 or 150–200 ms.

Third, after 200 ms the amount of information available in each 50-ms period starts to become lower, as shown in Fig. 12*A*. This is a strong indication that much of the information is available early on in the spike train, and that including information from later periods in the response has diminishing returns. The information available in these 50-ms response periods is reasonably correlated with the firing rate of the neurons, which tends to be highest early in the response, and to drop somewhat after this (e.g., see Fig. 3).

Fourth, it is shown in Fig. 12*B* that considerable information is present from the firing rate in periods as short as 20 ms (see, e.g., most of the 20-ms values between 100 and 220 ms). Indeed, the information available in the best 20-ms periods is 64.9% of that available from the whole period 100–500 ms shown in Fig. 12*A*. Further, it is of interest that some of these particular neurons in the temporal cortical visual areas start to respond as early or even just before 80 ms, so that useful information is available in the 80- to 100-ms epoch, and even some information in the 60- to 80-ms epoch from some of these neurons (see Fig. 8*B*).

The analyses presented in Fig. 12 imply that there might be diminishing returns in continuing to extract information from spike trains for longer and longer periods. We were able to assess this directly by performing an analysis of the

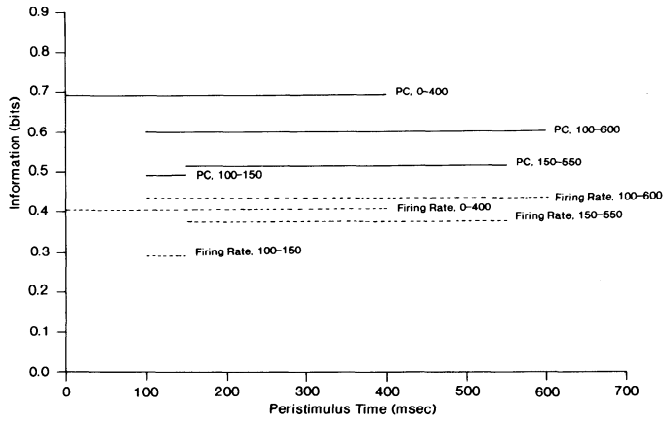

FIG. 9. Information that is available about which image was shown and where it was fixated from the 1st 3 principal components (——) and from the firing rate (- - -) for different poststimulus periods. *Time 0* is the onset of the visual stimulus. Correction 1 was applied. Data shown are the means over all 48 cells.

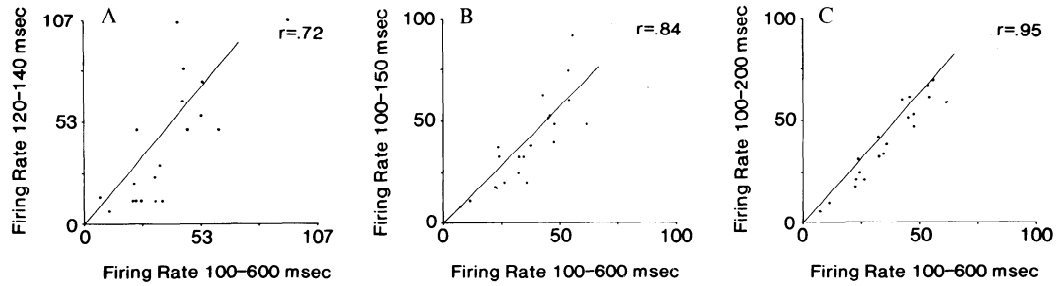

FIG. 10. Relation between the firing rates of a neuron measured over short periods (of 20, 50, and 100 ms) with the firing rate measured in a long period, of 500 ms. The correlation coefficient (r) is shown. Each data point represents the mean rate estimates for a particular image and a particular fixation position.

information available from the neuronal response starting at *time 0*, and extending on from this for different time periods, as shown in Fig. 13. It is clear that the rate of information extraction (the slope of the curve) is greatest in the period 100–200 ms, and that there are indeed diminishing returns in extending the analysis beyond this. This has important implications for understanding the period over which it would be useful for the next cortical stage in processing to attempt to accumulate information.

It is also possible to determine the amounts of information present in short time periods based on PCs1–3, which include information available in temporal encoding. Such an analysis performed for a 50-ms time period is shown in Fig. 9. (Data from 5 bins each 10 ms wide were included in this analysis.) In this case, the analysis was for the information about both image and position, so that possible latency of onset differences for different fixation positions could maximize what might be extracted by the principal component analysis. What is found is that indeed more information is available for the 100- to 150-ms period from PC than from firing rate analysis (see Fig. 9), but that even in this best case for PC analysis, the rate produces a reasonable proportion of all the information that can be extracted with the PC analysis.

Last, we assess whether for this population of cells information about the different aspects of the stimuli are reflected by the different principal components. The aspects of the stimuli that we were able to separate were image type (i.e., face identity, from a set of 4 faces) and fixation position on the screen (from a set of 5 positions). The proportion of the total information (i.e., about image and position) provided about the image is shown along the ordinates of Fig. 14, and about the position is plotted along the abscissae of Fig. 14. We show in Fig. 14 (PC1) the information available from principal component one. Each point represents the findings for one cell. The points (in comparison to the data for higher principal components, PC2 and PC3) show that the first principal component tends to provide more information about the image than about position (i.e., the majority of the points are above a 45° line superimposed on the graph). In contrast, in Fig. 14 (PC2 and PC3), the points tend to lie below a 45° line, indicating that these principal components provide more information about fixation position on the screen than about image type. (A statistical analysis based on the relative information about image vs. fixation position showed that the ratio image information/position information was significantly higher for PC1 than for PC2, Wilcoxon $P < 0.0001$.) In Fig. 14 (PC1–3), the information available from the first three principal components is shown. When all three principal components are included, the points are better distributed away from the axes of the graph, indicating that between them, the first three principal components provide information about both image type and fixation position. These

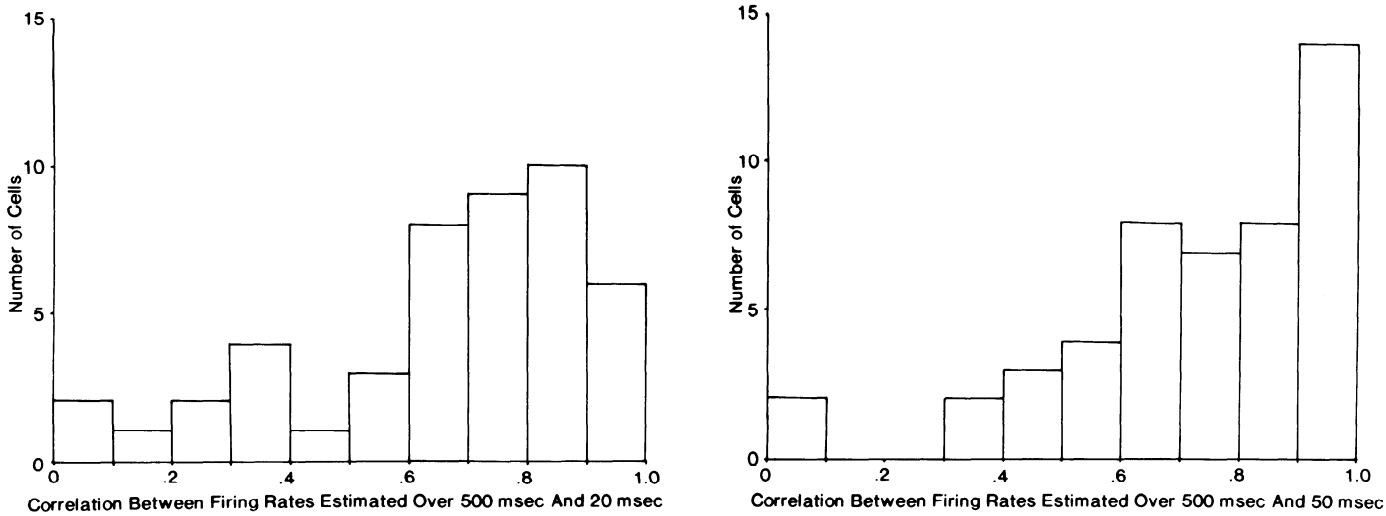

FIG. 11. Histograms showing for the population of cells analyzed the correlation between the firing rates estimated over 500 and 20 ms (*left*) or 50 ms (*right*).

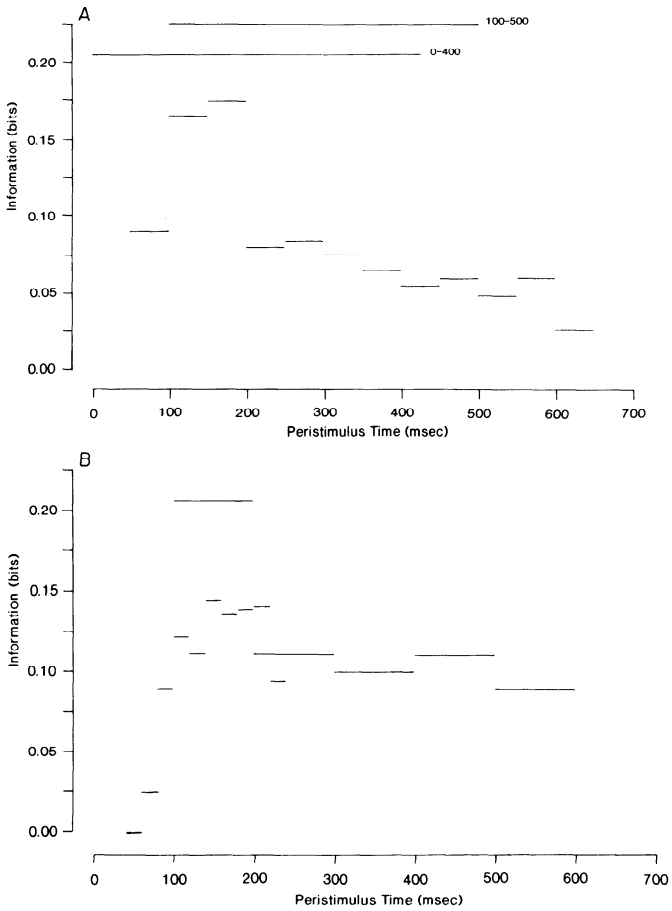

FIG. 12. *A*: information available about which image was shown from different (mainly 50 ms) time epochs of the neuronal activity. The length of the epoch, and the time at which it was taken with respect to stimulus onset, are indicated by the length and position of each horizontal line. Information shown reflects that obtained from the firing rates about which image was shown. Correction 1 was applied to the data. For comparison, the information available in 400-ms epochs starting at 0 and 100 ms is also shown. Data shown are the means over all 48 cells. *B*: as for *A*, except that the information available from 100- and 20-ms epochs is shown.

data are consistent with the evidence presented earlier that the image type is reflected in the firing rate of the cell, which is reflected by principal component 1, and that fixation position on the screen can produce different neuronal response latencies, which are reflected in the more time-varying higher principal components (see Fig. 4).

DISCUSSION

The results and analyses described here lead us to make the following points about temporal encoding in spike trains, principal component analysis of this, and the information available in short epochs of the response early in the response period of temporal cortical visual neurons.

First, the analyses indicate that of the first three principal components, the first provides by far the most information, with the second and third adding only small proportions (e.g., 18.8 and 8.4%, respectively, with the conservative correction 2; see Fig. 8). The relatively low amount of information provided by the second and higher principal components compared with some earlier analyses of Optican and Richmond (1987) is probably due in part to the use of a correction for limited sampling, the general form of which has been noted to be required by Optican et al. (1991). Our findings are thus consistent with those of Optican and Richmond (1987) and Optican et al. (1991) in that we confirm that there is information contained in more than the first principal component. However, our finding is that with the correction applied, relatively small proportions of information are provided by the second and third principal components (cf. Optican et al. 1991).

Eskandar et al. (1992) have very recently published data from inferior temporal cortex neurons responding to Walsh patterns and have again found that more information is contained with PC than with firing rate analysis. However, we are not clear whether correction 1, that advocated by Optican et al. (1991), was applied to the data obtained by Eskandar et al. (1992). In any case, it is worth noting that even the use of correction 1 will tend to overemphasize the contribution of PC analysis compared with firing rate analysis (see Fig. 8), relative to the more conservative correction 2.

Second, the magnitude of the second and higher principal components is even smaller if the spike train analysis is started after the onset of the neuronal response, instead of before the neuronal response has started (see Fig. 8*B*). Thus little temporal encoding about the stimulus is present (as shown by the small 2nd and higher principal components, and by the fact that the 1st is correlated with the mean firing rate) once the spike train has started. This further implies that some of the information available in the second and higher principal components reflects onset characteristics of the spike train. This is part of the evidence discussed below that suggests that variations in response latency are at least a part of what is reflected by the second and higher principal components.

Third, the first principal component appears to reflect the mean firing rate of the cell, as shown in Figs. 4 and 5, and as also made clear in Fig. 3. This is consistent with the

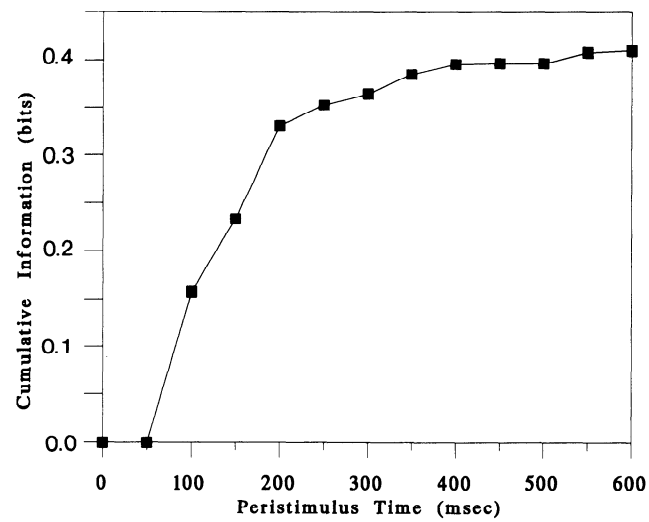

FIG. 13. Information available from the neuronal response starting at *time 0*, and extending on from this for different time periods. The information measure shown is the average over the 48 cells available from the firing rate about which image was shown and where it was fixated. Correction 1 was applied.

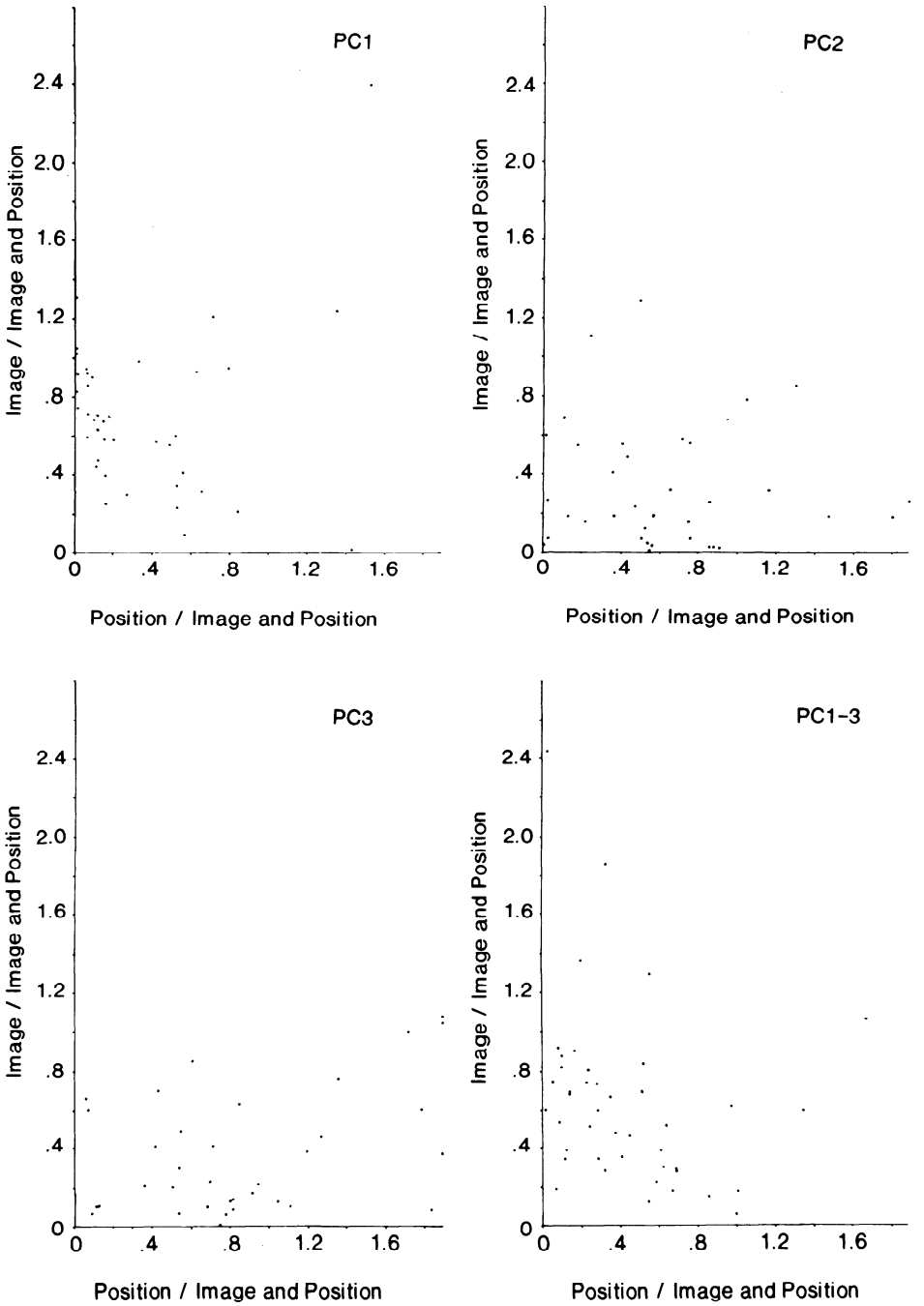

FIG. 14. Proportion of the total information available about the image is plotted on the ordinate, and about the fixation position is plotted on the abscissa. (The ratio can occasionally exceed 1 because of the bootstrap correction procedure used.) The separate graphs show this for each of the 1st 3 principal components separately, and for them combined. Each point represents data for 1 cell. Principal component 1 (PC1) tended to be related to image type more than fixation position, principal component 2 tended to be related to fixation position more than image type, and the combination of the 1st 3 principal components (PC1–3) tended to provide information both about the image and the fixation position.

finding of Richmond and Optican (1987). On the other hand, the second and higher principal components reflect at least partly the onset properties of the neuronal responses, such as response latency differences between the stimuli. Part of the evidence for this is noted under point two above. Further evidence comes from inspection of the waveforms of the second and higher principal components, which are such that appropriate weightings of these second and third principal components included in a reconstruction of the spike-density waveform from the principal components would produce differences in the onset latencies and characteristics to different stimuli. (Small differences in the responses latencies to the different stimuli were produced in the data from these temporal cortical neurons by the position of the stimulus on the retina. The variations in latency were small, of the order of 20 ms, and of position on the retina were large, of the order of 10°.)

Fourth, a considerable proportion of the information available from principal components 1–3 is available in the firing rate of the neuron, as shown in Fig. 9. This is particularly the case when the epochs are 150–550 ms, so that information about neuronal response latency is unlikely to be reflected in the principal component measure.

Fifth, periods of the firing rate much shorter than the 384-ms periods over which the information available with the use of temporal encoding has been calculated (Optican and Richmond 1987) are sufficient to give a reasonable estimate of the firing rate of the neuron. For example, in Figs. 10 and 11 it is shown that the firing rate estimated from sample periods of 50 ms is highly correlated with the

mean firing rate to a stimulus calculated over 500 ms. It is shown that, even if the sample period is as little as 20 ms, there is a reasonable correlation for most neurons with the firing rate calculated over the 500-ms period (Figs. 10 and 11).

Sixth, further direct evidence on the information available about the stimuli is shown by the analysis from information theory of the information available in short epochs shown in Fig. 12. Here it is shown that in sample periods of 50 ms the information available from the firing rate can be as high as 84.4% of that available from the firing rate calculated over 400 ms. (For comparison, with principal component analysis, the information about which image was shown available from a 50-ms period was found to be 77.1% of that available from the 0- to 400-ms period.) It is also shown in Fig. 12 that 71.2% of the information available from the firing rate over 400 ms is available in epochs as short as 20 ms.

Seventh, it is also clear that in this part of the brain most information is available in the early part of the neuronal response, as shown in Figs. 12 and 13 (cf. Eckhorn and Pöpel 1974, 1975). This is a further indication that taking long sample periods for the data to enable temporal encoding of information to be utilized brings diminishing returns as the sample period lengthens, although any added information would be expected to improve visual recognition.

Taken together, these analyses provide evidence that a short period of firing taken close to the start of the neuronal response provides a reasonable proportion of the total information that would be available if a long period of neuronal firing (e.g., 400 ms) were utilized to extract the information available from the firing rate and from temporal encoding. This finding thus is consistent with the hypothesis that, for each cortical area to provide useful information to the next, quite short periods of firing, of, for example, 20–50 ms, may be *sufficient.* This is in line with the evidence that, as visual signals are passed through the visual system, there may be a delay introduced by each stage of as little as 20–30 ms. This period may be sufficiently long to provide the information sufficient to support object recognition when visual stimuli are presented successively to the visual system (Thorpe and Imbert 1989). Analysis periods of 400 ms per cortical stage would perhaps allow each successive stage to extract the maximum amount of information available but would result in very long latencies for object recognition, if there are six cortical stages involved (Thorpe and Imbert 1989). In fact, recognition latencies are short, in the order of 300 ms. Further, we have now established that, in a visual backward masking paradigm in which a face is shown for 20 ms, face recognition can occur when the stimulus is followed 20 ms after its onset by a pattern mask, and that under these conditions neurons of the type described here respond for only 20–30 ms (E. T. Rolls, M. J. Tovée, D. G. Purcell, A. L. Stewart, and P. Azzopardi, unpublished observations; Tovée and Rolls, 1993). Of course, if more time is available, perception is better, and indeed under the conditions just described when the neurons fired in the masking paradigm for 20–30 ms, recognition by human observers was just possible.

A possibility might be that the firing that occurs early in the neuronal response of a particular cortical area might be nonselective, and a long period might be required before the firing reflected information about the stimulus in a reliable way that was also consistent with the neuron's later response. This is not the case, in that the firing rate in early response periods correlates well with that measured over a full 500-ms period (see Fig. 11). The same finding shows that the information reflected by the neuronal response is the same during the initial neuronal response as later on in the response period.

The implications for computation by the cerebral cortex are that quite short periods of the firing of a previous cortical area are sufficient for a particular cortical area to be provided with enough information to enable it to perform its computation; and for the actual computation to be performed so that information becomes available within 20–30 ms to the next cortical area. The fact that cortical neurons in the visual system have firing rates that are often as high as 100 Hz to effective visual stimuli is also an indication that the degree of activation of a cortical neuron can be assessed in a period as short as 20–50 ms. We note that, although an individual neuron might fire one to five spikes in this period, and this is sufficient to provide the amounts of information that we have quantified in this paper, it is the case that in the brain there would be many neurons that together would provide much larger amounts of information about the stimulus. One style of computation with which this is consistent is feed-forward processing whereby the neuronal output at a particular cortical stage is produced by applying the input firing as a vector of axonal firing to the synaptic weight vector of a cortical pyramidal cell, producing an activation that reflects the dot product of these two vectors (or possibly some more nonlinear function that involves multiplicative operations between active synapses close together on a dendrite), the elicitation of firing after a nonlinear operation such as a threshold, with then some time available for inhibitory interactions between nearby pyramidal cell that could implement competition through inhibitory interneurons (Rolls 1989, 1992). However, current theoretical investigations in which integrate and fire neurons with biophysically appropriate time constants are considered by Treves (1993) indicate that local recurrent feedback processing could contribute to the information reflected in neuronal responses in the initial 20–50 ms of the response period. On the other hand, for at least the initial recognition of visually presented objects, there may not be sufficient time for information to progress all the way to inferior temporal cortex (in which neurons have typical response latencies of 80–100 ms), and then back via back-projections to influence early visual cortical areas (e.g., V2, visual area 2), and then having influenced activity there, for this effect of feedback to influence cortical activity in the inferior temporal cortex in a way required for inferior temporal cortex to produce a useful output. Back-projections could be particularly important in cortical learning, and in recall and attention (Rolls 1989).

The analyses described here address the issue of the information available from the spike train of a single neuron. Of course, it is a possibility that information is available from the relative time of firing of different populations of neurons, as envisaged for example by von der Malsburg (1990), Eckhorn et al. (1988), and Engel et al. (1992).

However, we note that, for the primate inferior temporal visual cortex, oscillations that might reflect such temporal binding between subpopulations of neurons are not a prominent feature of the neuronal activity when static visual stimuli are detected (Tovée and Rolls 1992).

The findings described here were obtained from neuronal activity in the temporal cortical visual areas. The neurons analyzed were selective for faces. They provided a useful sample for investigation because neurons with such response properties can be found regularly in these cortical areas (Rolls 1992). We believe, because of the constraints on the processing time available for each cortical stage for recognition of most classes of visual object, and for stimuli in other modalities, that similar processing, in which there is insufficient time for temporal encoding to be utilized at least for rapid identification, will be a property of cortical processing common for object recognition irrespective of modality. However, we look forward to comparable analyses with other classes of visual stimuli, and in other modalities.

This research was supported by Medical Research Council Grant PG8513790 to E. T. Rolls.

Address reprint requests to E. T. Rolls.

Received 15 July 1992; accepted in final form 16 March 1993.

REFERENCES

BAYLIS, G. C., ROLLS, E. T., AND LEONARD, C. M. Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Res.* 342: 91–102, 1985.

BAYLIS, G. C., ROLLS, E. T. AND LEONARD, C. M. Functional subdivisions of temporal lobe neocortex. *J. Neurosci.* 7: 330–342, 1987.

ECKHORN, R., BAUER, R., JORDAN, W., BROSCH, M., KRUSE, W., MUNK, M., AND REITBAECK, H. J. Coherent oscillations: A mechanism for feature linking in the visual cortex? *Biol. Cybern.* 60: 121–130, 1988.

EKHORN, R. AND POPEL, B. Rigorous and extended application of information theory to the afferent visual system of the cat. I. Basic concepts. *Kybernetik* 16: 191–200, 1974.

EKHORN, R. AND POPEL, B. Rigorous and extended application of information theory to the afferent visual system of the cat. II. Experimental results. *Biol. Cybern.* 17: 7–17, 1975.

ENGEL, A. K., KONIG, P., KREITER, A. K., SCHILLEN, T. B., AND SINGER, W. Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends Neurosci.* 15: 218–226, 1992.

ESKANDAR, E. N., RICHMOND, B. J. AND OPTICAN, L. M. Role of inferior temporal neurons in visual memory. I. Temporal encoding of information about visual images, recalled images, and behavioral context. *J. Neurophysiol.* 68: 1277–1295, 1992.

FEIGENBAUM, J. D. AND ROLLS, E. T. Allocentric and egocentric spatial information processing in the hippocampal formation of the behaving primate. *Psychobiology* 19: 21–40, 1991.

FUKUNAGA, K. *Introduction to Statistical Pattern Recognition.* New York: Academic, 1972.

MACRAE, A. W. On calculating unbiased information measures. *Psychol. Bull.* 75: 270–277, 1971.

VON DER MALSBURG, C. A neural architecture for the representation of scenes. In: *Brain Organization and Memory: Cells, Systems and Circuits,* edited by J. L. McGaugh, N. M. Weinberger, and G. Lynch. New York: Oxford Univ. Press, 1990, p. 356–372.

MERRILL, E. G. AND AINSWORTH, A. Glass-coated platinum-plated tungsten microelectrodes. *Med. Biol. Eng.* 10: 662–672, 1972.

OPTICAN, L. 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: 132–146, 1987.

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.

RICHMOND, B. J. AND OPTICAN, L. 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 primary 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 characteristics. *J. Neurophysiol.* 57: 132–146, 1987.

ROLLS, E. T. Functions of neuronal networks in the hippocampus and neocortex in memory. In: *Neural Models of Plasticity: Experimental and Theoretical Approaches,* edited by J. H. Byrne and W. O. Berry. San Diego, CA: Academic, 1989, chapt. 13, p. 240–265.

ROLLS, E. T. Neural organisation of higher visual functions. *Curr. Opin. Neurobiol.* 1: 274–278, 1991.

ROLLS, E. T. Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 335: 11–21, 1992.

ROLLS, E. T., BURTON, M. J., AND MORA, F. Hypothalamic neuronal responses associated with the sight of food. *Brain Res.* 111: 53–66, 1976.

ROLLS, E. T., SANGHERA, M. K., AND ROPER-HALL, A. The latency of activation of neurons in the lateral hypothalamus and substantia innominata during feeding in the monkey. *Brain Res.* 164: 121–135, 1979.

ROLLS, E. T., YAXLEY, S., AND SIENKIEWICZ, Z. J. Gustatory responses of single neurons in the orbitofrontal cortex of the macaque monkey. *J. Neurophysiol.* 64: 1055–1066, 1990.

THORPE, S. J. AND IMBERT, M. Biological constraints on connectionist models. In: *Connectionism in Perspective,* edited by R. Pfeifer, Z. Schreter, and F. Fogelman-Soulie. Amsterdam: Elsevier, 1989, p. 63–92.

TOVÉE, M. J. AND ROLLS, E. T. Oscillatory activity is not evident in the primate temporal visual cortex with static stimuli. *Neuroreport* 3: 369–372, 1992.

TOVÉE, M. J. AND ROLLS, E. T. The neurophysiology of visual masking. *Brain Res. Assoc. Abstr.* 10:3, 1993.

TREVES, A. Mean field analysis of neuronal spike dynamics. *Network.* In press.