Object Perception in Natural Scenes: Encoding by Inferior Temporal Cortex Simultaneously Recorded Neurons

Nikolaos C. Aggelopoulos, Leonardo Franco, and Edmund T. Rolls
Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

Aggelopoulos, Nikolaos C., Leonardo Franco, and Edmund T. Rolls. Object perception in natural scenes: encoding by inferior temporal cortex simultaneously recorded neurons. *J Neurophysiol* 93: 1342–1357, 2005. First published October 20, 2004; doi:10.1152/ jn.00553.2004. The firing of inferior temporal cortex neurons is tuned to objects and faces, and in a complex scene, their receptive fields are reduced to become similar to the size of an object being fixated. These two properties may underlie how objects in scenes are encoded. An alternative hypothesis suggests that visual perception requires the binding of features of the visual target through spike synchrony in a neuronal assembly. To examine possible contributions of firing synchrony of inferior temporal neurons, we made simultaneous recordings of the activity of several neurons while macaques performed a visual discrimination task. The stimuli were presented in either plain or complex backgrounds. The encoding of information of neurons was analyzed using a decoding algorithm. Ninety-four percent to 99% of the total information was available in the firing rate spike counts, and the contribution of spike timing calculated as stimulus-dependent synchronization (SDS) added only 1–6% of information to the total that was independent of the spike counts in the complex background. Similar results were obtained in the plain background. The quantitatively small contribution of spike timing to the overall information available in spike patterns suggests that information encoding about which stimulus was shown by inferior temporal neurons is achieved mainly by rate coding. Furthermore, it was shown that there was little redundancy (6%) between the information provided by the spike counts of the simultaneously recorded neurons, making spike counts an efficient population code with a high encoding capacity.

INTRODUCTION

A fundamental issue is how information is encoded by populations of neurons (Franco et al. 2004; Gawne and Richmond 1993; Rolls and Deco 2002; Shadlen and Movshon 1999; Singer 1999, 2000; Treves 2000). It has been hypothesized that synchronization between sets of neurons might be used to indicate that the features represented by the different neurons should be grouped or bound together, thus facilitating segmentation of simultaneously present objects from each other and from the background. The hypothesis is that stimulus-dependent neuronal synchronization (SDS) would be useful, in that particular sets of features might need to be bound together for one object, but not for another object (Kayser et al. 2003; Malsburg 1990; Singer 1999; Singer and Gray 1995). We specifically address this hypothesis, and more generally, the relative quantitative contributions of SDS and firing rates to information encoding, by analyzing the responses of inferior temporal cortex neurons, where neurons respond to objects and faces (Desimone et al. 1984; Gross et al. 1972; Perrett et al. 1982, 1992; Rolls 2000; Rolls and Deco 2002; Rolls et al. 1994; Tanaka 1996). We used recently developed information theoretic methods (Franco et al. 2004; Rolls et al. 1997) to investigate information encoding when two objects simultaneously presented must be discriminated from each other and segmented from the background. Although SDS has been found in a number of test situations (Hatsopoulos et al. 1998; Singer 2000), it is important to know how significant a contribution it makes relative to the spike counts recorded from the neurons (Dan et al. 1998; Oram et al. 2001; Panzeri et al. 1999; Rolls et al. 2003b, 2004). Another fundamental issue is the extent to which neurons encode independent information versus whether redundancy is present (Gawne and Richmond 1993; Reich et al. 2001; Rolls et al. 2003b, 2004).

We applied information theoretic methods to the responses of neurons in the inferior temporal visual cortex recorded under conditions in which feature binding is likely to be needed; that is, when the monkey had to choose to touch one of two simultaneously presented objects, with the stimuli presented in a complex natural background. The investigation is thus directly relevant to whether SDS contributes to encoding under natural conditions. Neurons in the inferior temporal visual cortex respond in some cases to object features or parts and in other cases to whole objects, provided that the parts are in the correct spatial configuration (Desimone et al. 1984; Gross et al. 1972; Perrett et al. 1982, 1992; Rolls et al. 1994; Tanaka 1996; Vogels 1999), and so it is very appropriate to measure whether SDS contributes to information encoding in the inferior temporal visual cortex when two objects are present in the visual field and when they must be segmented from the background in a natural visual scene, which are the conditions in which it has been postulated that SDS would be useful (Kayser et al. 2003; Malsburg 1990; Singer 1999; Singer and Gray 1995).

METHODS

Recording techniques

The activity of single neurons was recorded with epoxy-insulated tungsten microelectrodes in a macaque monkey (*Macaca mulatta*; weight, ~8 kg) using techniques described previously (Booth and Rolls 1998). All procedures, including preparative and subsequent ones, were carried out in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were licensed under the UK Animals (Scientific Procedures) Act 1986. The action potentials of single neurons on several microelectrodes were amplified (Rolls et al. 1979) and viewed on-line during experiments.

Spikes from single neurons were isolated using Brainwave Enhanced Discovery data acquisition for off-line data analysis (DataWave), verifying as a final check that spikes of perfectly isolated neurons had been recorded by checking that no spikes occurred very close together in time in the interspike interval histogram. Eye position was monitored and measured with the scleral search coil technique (Judge et al. 1980) using 1-kHz digitization and storage of new values every 20 ms, with a calibration task performed during each recording session to provide an accuracy of better than 1°.

Task and stimuli

The monkeys performed a visual search task, in which on any trial, two images each subtending $9 \times 7°$ were presented simultaneously on a computer monitor, and the monkey could obtain two to three drops of fruit juice for every touch of the correct stimulus. On each trial, the monkey had to search for the position of the reward-related image and touch the image. A touch to the other stimulus resulted in the delivery of two to three drops of aversive saline. Different stimuli were used in different experiments. The monkey learned typically within five trials which was the reward stimulus of each pair, and data collection from the set on neurons in any one experiment started only after the monkey had learned which was the correct and which was the incorrect stimulus of each pair. This is thus a visual discrimination task that requires stimulus-reward learning. The monkeys' performance was >95% correct for the first touch. The two stimuli were placed with their centers 8.75° above and 8.75° below the center of the screen, with the position of the reward-associated and punishment-associated image randomized to above or below the screen center on every trial. The monitor was 23 cm away from the monkey. The whole screen subtended $55 \times 70°$ at the retina. The grayscale images were placed on either a blank background of mid-level gray (127/255) or on a complex natural scene, as shown in Fig. 1. The blank and complex backgrounds occurred in random order. The stimuli had a resolution of 64×64 pixels, but were prepared in such a way that they could be presented on either a complex background or a blank background that had a resolution of 512×512 pixels. The stimuli consisted of images of objects, faces, and geometrical patterns of the type that are effective in producing responses from inferior temporal cortex neurons (Rolls and Tovee 1995; Tamura and Tanaka 2001). The complex natural scene background was uniformly complex and did not allow easy segmentation in any particular region. If any of the neurons in an experiment responded to the normal background shown in Fig. 1, other comparable backgrounds were used, and in no experiment were different results found for different backgrounds.

In investigation 1, there were two stimulus pairs, and both stimuli in one simultaneously shown pair were selected to be effective for one or more of the neurons, and in the other pair, were selected to be ineffective for one or more of the neurons. As stated above, in each pair of images, one was rewarded, and the other was associated with punishment. It was possible to measure the information provided by the neurons that the image was the first pair of stimuli (the effective pair) versus the second pair of stimuli (the ineffective pair). This information was being measured in a task in which the monkey had to segment the two stimuli from their background and from each other, identify each stimulus, and decide which stimulus to touch. The hypothesis being tested in investigation 1 was that SDS might be present when one pair of the objects was being processed and might not be present, or might be present between different neuron pairs, when the second pair of objects was being presented; furthermore, how the information gained in this way compared with the information gained from the firing rates, which for some neurons were high to both stimuli of one simultaneously presented pair, but not the other simultaneously presented pair. (This SDS might, if present, be useful to the monkey in discriminating between the pair of stimuli presented simultaneously and might be present when the monkey was process-
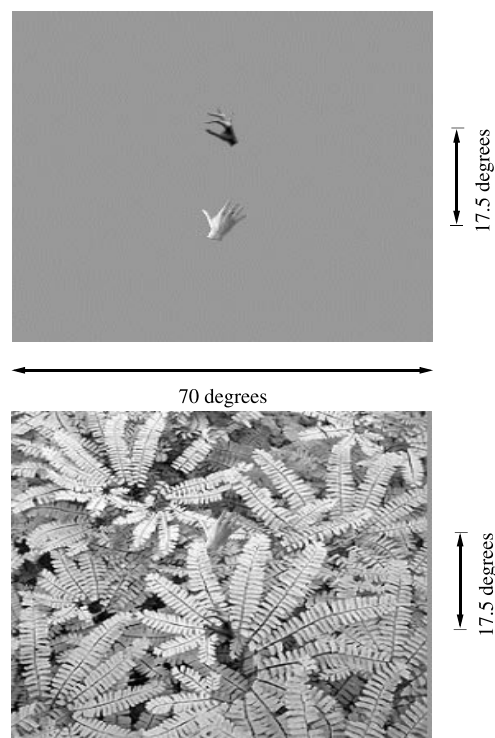

FIG. 1. Visual discrimination task. Two objects were presented on a screen subtending $70 \times 55°$ in either a blank background (*left*) or a complex natural scene (*right*). If the monkey touched 1 of the objects (S+), he obtained 2–3 drops of juice reward. If he touched the other object (the S−), he obtained saline. Positions of each object were randomized to the position above or below the screen center on each trial.

ing overtly or covertly one member of the pair of simultaneously presented stimuli.)

In investigation 2, there was one pair of stimuli, and one stimulus was selected to be effective for one or more of the neurons, and the other stimulus was selected to be ineffective for one or more of the neurons. In different experiments, either the effective stimulus, or the ineffective stimulus, was rewarded. [As shown previously, whether a stimulus was associated or not with reward in this and similar tasks does not influence the firing rate response of inferior temporal cortex neurons (Rolls et al. 1977, 2003a). In particular, Rolls et al.(2003a) showed that, provided that the monkey fixated a stimulus, the firing of inferior temporal cortex neurons was unaffected by whether it is a target being searched for or not, and in this sense, being attended to or not.] It was then possible to measure the information provided by the neurons that the stimulus was the effective stimulus at which the monkey was looking, or the ineffective stimulus. This was achieved by selecting 100-ms epochs of the firing rate in which the monkey's fovea was held still within the boundary of one or other of the stimuli from within a trial. This experiment thus enabled measurement of the firing rate of the neurons while the eyes moved from one stimulus to another dynamically during a trial in, for example, a complex natural scene, and how the information from the population was provided while the monkey was segmenting each stimulus from the background and identifying each stimulus, prior to deciding which stimulus to touch. An important part of the design was that, on every trial, the position of the objects was randomized for the upper or lower position, and the monkey had to find the correct position of the object that led to a juice reward when touched. The monkey normally fixated on the object he was about to touch, but before this, typically looked at both objects to determine where the object to touch was located (as is clearly shown in Fig. 7). On every trial in both investigations, the monkey took a decision in the first 300–400 ms about where to touch. Thus on every trial in this period, the feature binding and segmenta-

tion required in natural vision conditions is taking place, and this is the period in which we measured the information available from spike counts and SDS.

Thus vision for two objects in a complex scene is taking place in these experiments, and the design was useful for investigating the neuronal encoding required to implement visual object identification and selection.

Neurophysiological procedure

No more than six single neuron microelectrodes (tungsten insulated with epoxylite, 2–5 MOhm, FHC, Bowdinham, MA) were simultaneously lowered into the cortex in the superior temporal sulcus (STS) and the inferior temporal cortex (labeled T in Fig. 9 and defined in this study as the cortex forming the gyrus of the temporal lobe but lateral to the perirhinal cortex). The responses of isolated neurons were measured to a wide variety of small stimuli on the touchscreen. The isolation of neurons with these relatively high-impedance single neuron recording microelectrodes with tips of a few microns was good, in that these microelectrodes typically record from one and sometimes from two neurons, with signal to noise ratios that were typically >3:1. Indeed, for the majority of the recordings (78.2% of the 142 pairs of simultaneously recorded neurons in investigation 1 and for 74.7% of the 95 pairs in investigation 2), the neurons were obtained from different microelectrodes. The EPS system (Alpha-Omega) was used to move the electrodes independently until ≤6 neurons could be recorded at the same time. Stimuli used included faces, animals, and inanimate objects (for examples, see Rolls and Tovee 1995). For each experiment it was known that one to two of the typically three to five neurons did have different firing rates to the stimuli (quite typical for inferior temporal cortex neurons, which are likely to respond at >50 spikes/s for the most effective stimulus in a set), and these one to two neurons, and the other neurons, could have had stimulus-dependent correlations (although this was not evaluated at the time of the recording). Thus the sets of simultaneously recorded neurons were not especially selected to have rate versus synchrony-related information. Indeed, in both investigations 1 and 2, some of the simultaneously recorded neurons did not have significant firing rate responses or significantly different firing rate responses to the two stimuli, so there was plenty of opportunity for neurons with no rate information to contribute to the information available. It was also a condition for running the experiment that the neurons did not respond to the background. Most anterior inferior temporal cortex neurons at coordinates that were typically 2–7 mm posterior to the sphenoid reference (see Fig. 9) did not respond to the background image, which for all experiments was as shown in Fig. 1.

Neuronal responses were recorded during a sequence of 160–240 presentations (trials). The targets occurred randomly in one of two positions on the screen, either 8.75° above or 8.75° below the center of the screen. This was intended to minimize search time so that the main task would be one of segregating the two stimuli. Each trial was preceded by a 0.5 s tone cue to enable the monkey to look toward the center of the screen before the visual display appeared. The monkey was allowed to touch up to four times to obtain separate aliquots of fruit juice reward before the next trial. Trials in which the target object appeared in a blank screen or a natural scene were run in random order.

Recording sites

X-rays were taken at the end of each recording session to determine the position of the microelectrode, relative to bony landmarks and the permanently implanted reference electrodes. At the end of the final tracks, microlesions were made in the areas of cortex in which recordings were made to mark typical recording sites (Feigenbaum and Rolls 1991). Reconstructions of the tracks were made in serial 50-μm histological sections using the positions of the microlesions,

the reference electrodes in the histology, the corresponding X-ray coordinates, and the X-ray coordinates of all recorded cells to determine the locations of all the cells.

Data analysis

The aim of the data analysis was to obtain measures of the information in the firing of the neurons and to separate the information contained in the firing rates to that contained in the relative timing of the spikes across neurons. We applied recently developed information theoretic techniques to quantify these contributions (rate and the relative timing of spikes from different cells) that use a decoding method that can operate with large numbers of neurons and spikes from each neuron (Franco et al. 2004; Rolls et al. 1997). Spikes over the period starting 100 ms after stimulus onset (the typical latency for the neuronal responses) were included in the analyses, with the epochs described in the description of the two investigations. Data collected were analyzed separately depending on whether the stimuli appeared against the plain and complex backgrounds.

Information measurement algorithm

The direct approach to compute the information about a set of stimuli conveyed by the responses of a set of neurons is to apply the Shannon mutual information measure (Cover and Thomas 1991; Shannon 1948)

$$I(s, \vec{r}) = \sum_{s \in S} \sum_{\vec{r}} P(s, \vec{r}) \log_2 \frac{P(s, \vec{r})}{P(s)P(\vec{r})} \quad (1)$$

where $P(s, \vec{r})$ is a probability table embodying a relationship between the variable s (here, the stimulus) and $\vec{r}$ (a vector where each element is the firing rate of 1 neuron).

However, because the probability table of the relation between the neuronal responses and the stimuli, $P(s, \vec{r})$, is so large [given that there may be many stimuli and that the response space is very large, growing exponentially with the number of neurons for the rate information (Panzeri et al. 1999; Treves and Panzeri 1995), and even more if relative spike timing is considered], in practice, it is difficult to obtain a sufficient number of trials for every stimulus to generate the probability table accurately, at least with data from mammals, in which the experiment cannot usually be continued for many hours of recording from a whole population of cells. To circumvent this undersampling problem, Rolls et al. (1997) developed a decoding procedure, in which an estimate (or guess) of which stimulus (called s') was shown on a given trial is made from a comparison of the neuronal responses on that trial with the responses made to the whole set of stimuli on other trials. One then obtains a conjoint probability table $P(s, s')$, and the mutual information I_p based on probability estimation (PE) decoding between the estimated stimulus s' and the actual stimulus s that was shown can be measured

$$I_p = \sum_{s \in S} \sum_{s' \in S} P(s, s') \log_2 \frac{P(s, s')}{P(s)P(s')} = \quad (2)$$

$$= \sum_{s \in S} P(s) \sum_{s' \in S} P(s'|s) \log_2 \frac{P(s'|s)}{P(s')} \quad (3)$$

These measurements are in the low-dimensional space of the number of stimuli, and therefore the number of trials of data needed for each stimulus is of the order of the number of stimuli, which is feasible in experiments. In practice, it is found that for accurate information estimates with the decoding approach, the number of trials for each stimulus should be at least twice the number of stimuli (with a minimum of 16 trials for each stimulus) (Franco et al. 2004). The advantage of the decoding method (Franco et al. 2004) used here over earlier methods that directly compute the Shannon information

(Hatsopoulos et al. 1998; Oram et al. 2001; Panzeri et al. 1999; Rolls et al. 2003b, 2004), is that the decoding method works successfully with large numbers of simultaneously recorded neurons and with large numbers of spikes from each neuron. The direct methods (Panzeri et al. 1999; Rolls et al. 2003b, 2004), even with few stimuli, need many more trials than are available here if the information is to be measured from more than very short epochs consisting essentially of one spike from each neuron, because the probability space between each stimulus and the response measures for every neuron becomes so large (larger than the number of stimuli) (Rolls et al. 1997; Treves and Panzeri 1995). It is for this reason that we used the decoding approach described here, knowing also that it measures information that is close to what could be measured directly, as shown by Franco et al. (2004).

The decoding procedure essentially compares the vector of responses on a single trial with the average response vectors obtained previously to each stimulus. This decoding can be as simple as measuring the correlation, or dot (inner) product, between the test trial vector of responses and the response vectors to each of the stimuli. In this paper, we used a Bayesian procedure based on a Gaussian assumption of the spike probability distributions as described in detail by Rolls et al. (1997, 2003b). The new step introduced by Rolls et al. (2004) and used in this paper is to introduce into the Table Data $(s, \vec{r})$ new columns containing a measure of the cross-correlation (averaged across trials) for some pairs of cells (see example in Fig. 2C). The decoding procedure can take account of any cross-correlations between pairs of cells and thus measure any contributions to the information from the population of cells that arise from cross-correlations between the neuronal responses. If these cross-correlations are stimulus-dependent, their positive contribution to the information encoded can be measured. We note that the information measured with any decoding procedure provides a lower bound on the true information that might be measured directly but that the decoding
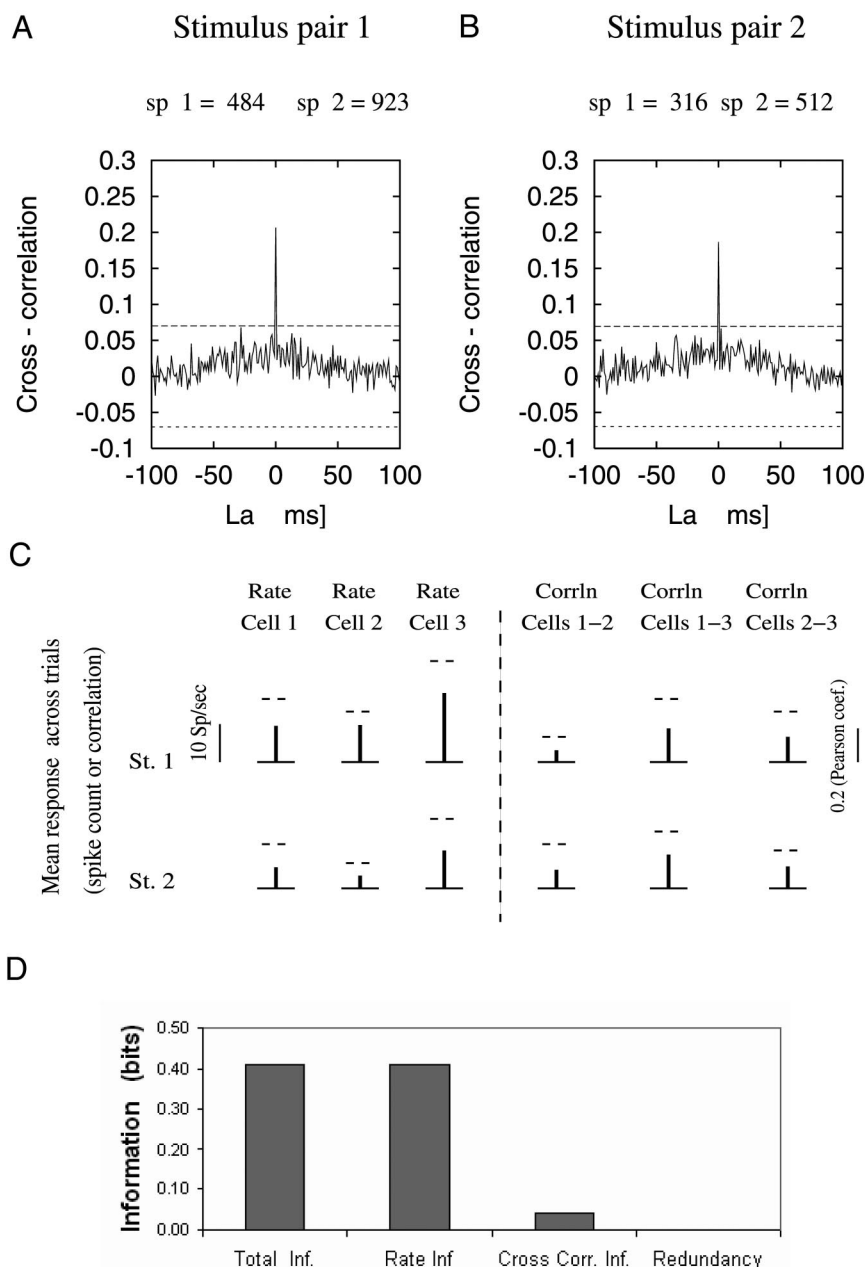
Blank background

FIG. 2. Example of a set of neurons with cross-correlations that are not stimulus-dependent recorded in investigation 1 in a blank background. *A* and *B*: cross-correlogram for a pair of neurons from experiment bj287 for 2 different stimulus pairs (1 and 2). Peak in the cross-correlogram located at lag 0 is present in both conditions and so is not stimulus-dependent. Number of spikes from each neuron used to construct the cross-correlograms is shown. Dashed horizontal lines show the 95% CI of the cross-correlation estimate. *C*: *left*: average firing rate (or equivalently spike count) responses of each of 3 cells (labeled as Rate Cell 1,2,3) to a set of 2 stimuli, St 1 and St 2. *Right*: measure of the cross-correlation (averaged across trials) for some pairs of cells (labeled as Corrln Cells 1–2 etc.). Variability of responses is indicated by the horizontal dashed line showing the SD of the mean calculated across the 40 trials available. From the responses on a single trial, probability $P(s')$ obtained by decoding the stimulus s was computed, based on values of both rates and cross-correlations. *D*: total information available from both rate and correlations, rate information (obtained from the number of spikes), information from any stimulus-dependent cross-correlations, and rate covariation redundancy between the neurons, is shown. Analysis epoch was 400 ms (experiment bj287).

procedure has been validated and shown to be efficient by Franco et al. (2004).

Further details of the decoding procedures (which have been validated by Franco et al. 2004) are as follows. The full probability table estimator (PE) algorithm uses a Bayesian approach to extract $P(s'|\vec{r})$, for every single trial from an estimate of the probability $P(\vec{r}|s')$ of a stimulus-response pair made from all the other trials (as shown in Bayes' rule—*Eq. 4*) in a cross-validation procedure

$$P(s'|\vec{r}) = \frac{P(\vec{r}|s')P(s')}{P(\vec{r})} \qquad (4)$$

where $P(\vec{r})$ (the probability for the vector $\vec{r}$ containing the firing rate of each neuron) is obtained as

$$P(\vec{r}) = \sum_{s'} P(\vec{r}|s')P(s') \qquad (5)$$

This requires knowledge of the response probabilities $P(\vec{r}|s')$, which can be estimated for this purpose from $P(\vec{r}, s')$, which is equal to $P(s')\prod_c P(r_c|s')$, where r_c is the firing rate of cell c. We note that $P(r_c|s')$ is derived from the responses of cell c from all of the trials except for the current trial, for which the probability estimate is being made. The probabilities $P(\vec{r}, s')$ are fitted with a Gaussian distribution whose amplitude at r_c gives $P(r_c|s')$. By summing over different test trial responses to the same stimulus s, we can extract the probability, that by presenting stimulus s, the neuronal response is interpreted as having been elicited by stimulus s'

$$P(s'|s) = \sum_{r \in \text{test}} P(s'|\vec{r})P(\vec{r}|s) \qquad (6)$$

After the decoding procedure, the estimated relative probabilities (normalized to 1) were averaged over all "test" trials for all stimuli to generate a (regularized) table $P_N^R(s,s')$ describing the relative probability of each pair of actual stimulus s and posited stimulus s' (computed with N trials). From this probability table, the mutual information measure I_p was calculated as described in *Eq. 3*. We note that any decoding procedure can be used in conjunction with information estimates both from the full probability table (to produce I_p) and from the most likely estimated stimulus for each trial in a frequency table $P_N^F(s,s)$ (to produce I_{ml}).

Because the probability tables from which the information is calculated may be unregularized with a small number of trials, a bias correction procedure to correct for the undersampling is applied (Panzeri and Treves 1996; Rolls et al. 1997). The correction term, C_1, to be used takes the form

$$C_1 \approx \frac{1}{2N\log(2)} \sum_s P(s) \sum_{s'} \left[\frac{Q_N^R(s,s')}{P_N^R(s,s')} - \frac{P_N^R(s,s')}{P(s)} \right]$$
$$- \frac{1}{2N\log(2)} \sum_{s'} \left[\frac{Q_N^R(s')}{P_N^R(s')} - P_N^R(s') \right] \qquad (7)$$

where $Q_N^R(s,s')$ is the table obtained analogously to $P_N^R(s,s')$, but averaging over all test trials $P^2(s'|r)$ instead of $P(s'|r)$, and where care has to be taken in performing the sums over s', to avoid including stimuli posited to have zero probability. For a derivation of this and other correction terms, and for that required to correct $I(s, s^P)$, we refer to Panzeri and Treves (1996). In practice, the bias correction that is needed with information estimates using the decoding procedures described here and by Rolls et al. (1997) is small, typically <10% of the uncorrected estimate of the information, provided that the number of trials for each stimulus is in the order of twice the number of stimuli (with a minimum of 16 trials for each stimulus).

We note that if Bayesian decoding is used, an assumption is that the joint probability distribution of the spike count responses of the cells is approximated by the product of the separate probability distribu-

tions for each cell. This approximation holds if the distributions are independent and may be less exact if there are correlations between the neurons' responses. In practice, this is not a limitation of the method in that the level of correlations found in practice produce only a relatively small distortion of the probability values used to compute the information, partly because these probability values are normalized before being used, reducing the distortion especially when relatively few (e.g., 40) trials of data per stimulus are used.

The data from the neuronal activity used to compute the joint probability distribution $P_N^R(s,s')$ was as follows. From the response of each cell c to each stimulus, we extracted a single mean spike count in a fixed time window (or firing rate, r_c, expressed in spikes per second).

The measure of the cross-correlation that was introduced into the Table Data $(s, \vec{r})$ on each trial was the value of the Pearson cross-correlation coefficient calculated for that trial at the appropriate lag for cell pairs that had significant cross-correlations. This value of this Pearson cross-correlation coefficient for a single trial was calculated from pairs of spike trains on a single trial by forming for each cell a vector of 0s and 1s, the 1s representing the time of occurrence of spikes with a temporal resolution of 1 ms. Resulting values within the range -1 to 1 were shifted to obtained positive values. An advantage of the Pearson cross-correlation coefficient is that it measures the amount of synchronization between pairs of neurons independently of the firing rate of the neurons. The lag at which the cross-correlation measure was computed for every single trial, and whether there was a significant cross-correlation between neuron pairs, was identified from the location of the peak in the cross-correlogram taken across all trials. (In all 28 significant cross-correlations of the 284 tested in investigation 1, all 28 were located at a lag of 0 ms, and the same was the case in investigation 2.) The cross-correlogram was calculated by, for every spike that occurred in one neuron, incrementing the bins of a histogram that corresponded to the lag times of each of the spikes that occurred for the other neuron, with a precision of ±1 ms. (This 3-ms bin width was sufficient to encompass the width of the cross-correlations found in the neurons described in this paper. Furthermore, we confirmed that extending the bin width to 7 ms did not increase the SDS-related information.) The raw cross-correlogram was corrected by subtracting the "shift predictor" cross-correlogram (which was produced by random re-pairings of the trials) to produce the corrected cross-correlogram. It was normalized to be in the range ±1. When calculating the stimulus-dependent cross-correlation information, we followed the procedure described by Franco et al. (2004) of including subtraction of any chance contribution to the stimulus-dependent correlation information using trial shuffling within a stimulus. The values of the correlations between the spike timings measured on every trial were shown (Franco et al. 2004; Hatsopoulos et al. 1998) to follow an approximately Poisson distribution, as did the firing rate counts, and the decoding algorithm used here has been shown to operate efficiently with such data (Franco et al. 2004). The decoding was performed by a truncated Gaussian fit to the data values obtained, because this has one more parameter than a Poisson fit and so can be more accurate, especially because the firing rate counts are distributed with slightly more variability than would be predicted from a Poisson distribution (see paragraph on the Fano factor in RESULTS). Full details and validation are provided by Franco et al. (2004).

We estimated the redundancy in the rate information by shuffling the order of the trials within a stimulus and comparing this to the measured rate information. We use the term "rate covariation redundancy" for this in this paper, because the term captures the extent to which the firing rate responses of the neurons covary within a trial and interact with the similarity of the average response profiles of the neurons to the set of stimuli (see Franco et al. 2004 for details of this term, also referred to as the stimulus-independent rate information in Oram et al. 1998, and Rolls et al. (2003b, 2004) for further discussion of the underlying concepts).

RESULTS

Investigation 1

It was possible to complete 31 experiments in which with 2–4 electrodes, 2–5 neurons were simultaneously recorded in the inferior temporal visual cortex for >40 trials while the monkey performed the visual discrimination task, touching the screen on every trial to obtain rewards if the correct image of the two being shown on the screen was touched. The total number of neurons in the sample was 109. All neurons recorded in any one experiment that had significant differences of firing rates to the stimuli, or significant cross-correlations, were included in the analysis. In this experiment, there were two stimulus pairs, as shown in Fig. 1, and both stimuli in one simultaneously shown pair were selected to be effective for one or more of the neurons, and in the other pair, were selected to be ineffective for the one or more of the neurons. It is emphasized that, on each trial, the monkey had to discriminate between the two stimuli being shown, in that a touch to one was rewarded and to the other was punished, so that the test conditions are directly relevant to testing the hypotheses in the Introduction that binding between features of an object and to perform segmentation from the background might be imple-

mented by SDS. The trials with a plain or complex background (see Fig. 1) were shown in random sequence.

The results for an experiment (bj287) in which cross-correlograms were present between some of the neuron pairs, but were not stimulus-selective, are shown in Fig. 2. Figure 2, *A* and *B,* shows the cross-correlations for one pair of neurons in the blank background. The cross-correlations were measured over an epoch of 400 ms starting 100 ms after stimulus onset. The cross-correlation was located at 0 ms and was significant. (The dashed horizontal lines show the 95% CI of the cross-correlation estimate.) Figure 2*C* shows the average firing rates of each of the neurons to each of the stimuli, and at the *right* of the diagram, the average cross-correlation values from the three pairs of neurons with the highest values. It is clear that at least some of the neurons had different firing rates to the two stimuli. The decoding algorithm (Bayesian full probability estimation) was applied to the data to estimate for each trial which stimulus (s') was shown by comparison with the data from all the other trials (which have values close to those shown in Fig. 2*C*, which is the average response over all trials). The results of calculating the information $I(s, s')$ in a 400-ms epoch from the spike counts only was 0.41 bits, from the

TABLE 1. *Experiment 1: information available in 400 ms about which stimulus was shown in experiment 1*

	Rate Info	Stimulus-dep Correlation Info	Rate Covariation Redundancy	Total Info	Rate Info	Stimulus-dep Correlation Info	Rate Covariation Redundancy	Total Info
	Plain	Plain	Plain	Plain	Complex	Complex	Complex	Complex
bj272	0.22	0.02	0.03	0.22	0.03	0.00	0.00	0.03
bj274b	0.11	0.00	−0.02	0.11	0.23	0.00	−0.04	0.23
bj274c	0.17	0.00	0.01	0.17	0.11	0.00	0.00	0.11
bj275a	0.18	0.00	−0.02	0.18	0.18	0.00	−0.01	0.18
bj275b	0.60	0.04	0.00	0.60	0.30	0.04	−0.02	0.30
bj277	0.86	0.00	−0.07	0.86	0.74	0.03	−0.11	0.74
bj279	0.42	0.02	−0.02	0.44	0.16	0.00	−0.02	0.16
bj280a1	0.84	0.00	−0.04	0.84	0.69	0.03	0.00	0.69
bj280a2	0.55	0.01	0.05	0.56	0.43	0.01	−0.05	0.44
bj280b	0.44	0.01	0.00	0.44	0.29	0.00	0.04	0.29
bj283	0.07	0.03	0.01	0.13	0.21	0.02	0.01	0.23
bj284	0.08	0.00	0.00	0.08	0.01	0.00	0.00	0.01
bj285a	0.73	0.03	−0.01	0.73	0.76	0.00	−0.04	0.76
bj285b	0.56	0.06	−0.05	0.56	0.26	0.00	0.00	0.26
bj287	0.41	0.04	0.00	0.41	0.21	0.00	−0.05	0.22
bj288	0.59	0.01	0.03	0.59	0.41	0.01	−0.06	0.41
bj289a	0.48	0.04	−0.08	0.52	0.05	0.00	0.01	0.05
bj289b	0.28	0.00	−0.06	0.28	0.01	0.00	0.00	0.01
bj292	0.22	0.00	0.01	0.22	0.08	0.00	0.00	0.08
bj293a	0.66	0.00	0.02	0.66	0.51	0.00	−0.04	0.51
bj293b	0.76	0.00	−0.01	0.77	0.40	0.03	−0.05	0.43
bj297	0.61	0.00	0.01	0.61	0.38	0.00	0.02	0.38
bj298a	0.56	0.04	−0.06	0.58	0.12	0.04	0.02	0.13
bj298b	0.42	0.05	0.01	0.42	0.01	0.00	0.00	0.01
bj320	0.65	0.01	0.00	0.65	0.59	0.02	−0.06	0.59
bj365a	0.61	0.05	−0.04	0.61	0.02	0.00	0.00	0.02
bj365b	0.65	0.03	−0.09	0.66	0.34	0.01	0.00	0.35
bj372	0.38	0.00	−0.03	0.38	0.60	0.02	−0.07	0.60
bj387a	0.16	0.00	−0.02	0.16	0.12	0.00	0.00	0.12
bj390	0.41	0.05	−0.02	0.41	0.09	0.01	−0.01	0.11
bj397	0.26	0.05	0.02	0.29	0.08	0.02	−0.02	0.10
Averages	0.449	0.018	−0.014	0.455	0.272	0.009	−0.018	0.275
Percentage of total	98.7%	4.0%	−3.1%	100%	98.9%	3.3%	−6.5%	100%

The contributions in bits of different components of the information extracted using a decoding algorithm. The information available in a 400-ms time window starting 100 ms after the stimulus onset for the cases of complex and plain backgrounds from 31 sets of 2–4 simultaneously recorded inferior temporal cortex neurons when presented with a pair of stimuli.

TABLE 2. *Experiment 1: information available in 100 ms about which stimulus was shown in experiment 1*

	Rate Info	Stimulus-dep Correlation Info	Rate Covariation Redundancy	Total Info	Rate Info	Stimulus-dep Correlation Info	Rate Covariation Redundancy	Total Info
	Plain	Plain	Plain	Plain	Complex	Complex	Complex	Complex
bj272	0.07	0.01	0.00	0.07	0.00	0.00	0.00	0.00
bj274b	0.01	0.00	0.00	0.01	0.13	0.02	−0.02	0.13
bj274c	0.14	0.01	0.00	0.14	0.02	0.00	−0.01	0.02
bj275a	0.06	0.00	−0.01	0.06	0.17	0.01	−0.01	0.17
bj275b	0.26	0.02	0.00	0.28	0.15	0.01	0.00	0.15
bj277	0.51	0.06	0.00	0.52	0.23	0.00	−0.07	0.23
bj279	0.25	0.02	0.04	0.27	0.03	0.01	−0.01	0.03
bj280a1	0.39	0.00	0.04	0.39	0.45	0.00	0.03	0.45
bj280a2	0.10	0.00	−0.03	0.10	0.06	0.00	−0.01	0.06
bj280b	0.20	0.00	−0.04	0.20	0.27	0.00	−0.03	0.27
bj283	0.02	0.01	0.00	0.04	0.02	0.01	0.00	0.02
bj284	0.02	0.00	0.00	0.02	0.04	0.01	−0.02	0.05
bj285a	0.28	0.00	0.01	0.28	0.38	0.00	−0.01	0.38
bj285b	0.17	0.04	0.00	0.20	0.02	0.01	0.00	0.03
bj287	0.09	0.00	−0.01	0.09	0.06	0.00	0.01	0.06
bj288	0.40	0.00	0.03	0.40	0.09	0.00	0.02	0.09
bj289a	0.13	0.01	0.02	0.13	0.01	0.00	0.00	0.01
bj289b	0.05	0.00	0.00	0.05	0.01	0.02	0.00	0.03
bj292	0.02	0.00	0.00	0.02	0.00	0.00	0.00	0.00
bj293a	0.17	0.00	−0.05	0.17	0.12	0.00	−0.01	0.12
bj293b	0.12	0.00	−0.01	0.12	0.06	0.00	0.01	0.06
bj297	0.32	0.02	0.03	0.33	0.10	0.00	0.01	0.10
bj298a	0.12	0.00	0.00	0.12	0.00	0.00	0.00	0.00
bj298b	0.15	0.00	0.04	0.15	0.00	0.00	0.00	0.00
bj320	0.17	0.00	0.00	0.17	0.26	0.04	−0.04	0.26
bj365a	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00
bj365b	0.28	0.04	−0.01	0.28	0.09	0.00	−0.03	0.09
bj372	0.33	0.00	0.00	0.33	0.27	0.00	−0.01	0.27
bj387a	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00
bj390	0.16	0.01	0.01	0.16	0.01	0.00	0.00	0.01
bj397	0.15	0.01	0.00	0.16	0.04	0.00	0.00	0.04
Averages	0.166	0.008	0.001	0.170	0.099	0.005	−0.006	0.100
Percentage of total	97.6%	4.7%	0.6%	100%	99.0%	5.0%	−6%	100%

The contributions in bits of different components of the information extracted using a decoding algorithm. The information available in a 100-ms time window starting 100 ms after the stimulus onset for the cases of complex and plain backgrounds from 31 sets of 2–4 simultaneously recorded inferior temporal cortex neurons when presented with a pair of stimuli.

cross-correlations only was 0.04 bits, and with the total information using both spike counts and cross-correlation information was 0.41 bits, as shown in Fig. 2*D* and Table 1. Thus, in the blank background, most of the information was available in the spike counts, with much less in the cross-correlations. In addition, Fig. 2*D* shows that the rate covariation redundancy in the spike counts across neurons (related to any similarity of the firing rate tuning profiles of the set of neurons to the different stimuli and the trial-by-trial covariation of the rates of the different neurons) was very low (0.0 bits). (Negative information in the rate covariation column of Tables 1 and 2 indicates redundancy, that is, that there is less information with simultaneously recorded neurons because of covariations of the firing rates of the different neurons on a trial by trial basis.)

The corresponding data from the same experiment (bj287) performed in the complex natural background are shown in Fig. 3. The cross-correlations between the same pair of neurons were similar in the complex (Fig. 3, *A* and *B*) and plain (Fig. 2, *A* and *B*) backgrounds, and the firing rates were also comparable (Fig. 3*C* vs. 2*C*) in the two backgrounds. The results of calculating the information $I(s, s')$ in the complex background from the spike counts only was 0.21 bits, from the cross-

correlations only was 0.0 bits, and with the total information using both spike counts and cross-correlation information was 0.21 bits, as shown in Fig. 3*D* and Table 2. Thus also in the complex background, most of the information was available in the spike counts, with much less in the cross-correlations. In addition, Fig. 3*D* shows that the rate covariation redundancy was higher (−0.05 bits). Comparison of the analyses summarized in Figs. 2 and 3 show that, in this experiment, less information was available in the complex background, and that this was due to the greater variability of the neuronal response (in particular in this experiment of cell 2) in the complex background than in the plain background.

The results of comparable analyses on a set of neurons (in experiment bj293b) with clear stimulus-dependent cross-correlations between some of the neuron pairs are shown in Figs. 4 and 5. Figure 4, *A* and *B,* shows that the cross-correlation between cells 1 and 2 was not very significant in the blank background but may have been greater for stimulus pair 1 than stimulus pair 2. The firing rates of the cells to the different stimuli were clearly different (Fig. 4*C*). The results of calculating the information $I(s, s')$ in a 400-ms epoch in the plain background from the spike counts only was 0.76 bits, from the

Complex background

A Stimulus pair 1

sp 1 = 520   sp 2 = 937

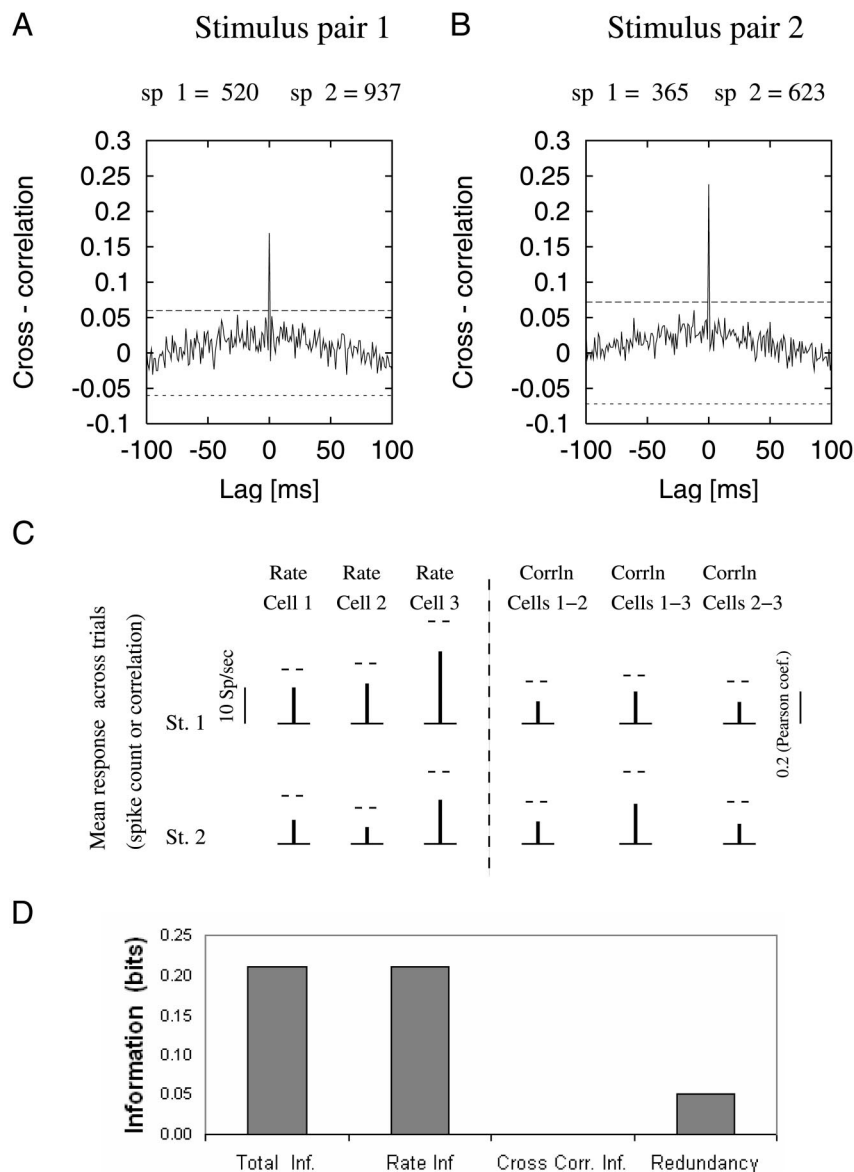B Stimulus pair 2

sp 1 = 365   sp 2 = 623

C

D

FIG. 3. Example of a set of neurons recorded with cross-correlations that are not stimulus-dependent in investigation 1 in a complex background. Conventions as in Fig. 2 (experiment bj287).

cross-correlations only was 0.0 bits, and with the total information using both spike counts and cross-correlation information was 0.77 bits (Fig. 4*D*).

Figure 5, *A* and *B,* shows that the cross-correlation in the same experiment between cells 1 and 2 was strongly stimulus-dependent, with a high and very significant cross-correlation only for stimulus pair 1. The firing rates of the cells to the different stimuli were clearly different (Fig. 5*C*). The results of calculating the information $I(s, s')$ in the complex background from the spike counts only was 0.40 bits, from the cross-correlations only was 0.03 bits, and with the total information using both spike counts and cross-correlation information was 0.43 bits (Fig. 5*D*). It is of interest that even though this set of neurons showed clear stimulus-dependent cross-correlations between some of its members, the amount of information they provided was quite low (0.03 bits) relative to that provided by the spike counts (0.41 bits).

Table 1 and Fig. 6 (*top*) summarize the data across all experiments in investigation 1 with a 400-ms analysis epoch, shown separately for plain and complex backgrounds. First, it is clear that, on average across the 31 experiments, the information related to the firing rate (0.449 bits) was much greater than the stimulus-dependent cross-correlation information (0.018 bits; for the plain background). This difference is also evident in the complex background (average rate information across experiments = 0.272 bits and average stimulus-dependent cross-correlation information = 0.009 bits). Second, in the plain background, the rate covariation redundancy was quite low (−0.014 bits compared with the rate information of 0.449 bits). In the complex background, the average rate covariation redundancy was a little higher (−0.018 bits compared with the rate information of 0.272 bits). This reflects greater redundancy in the complex background (6.6 vs. 3.1% in the plain background), which could arise not only because the tuning profiles to the stimuli become more similar in the complex background, but also perhaps because of any minor common response of the neurons to the background stimulus itself. Third, we note that the rate information shown in Table 1 in

Blank background

A   Stimulus pair 1

sp 1 = 118   sp 2 = 174

B   Stimulus pair 2
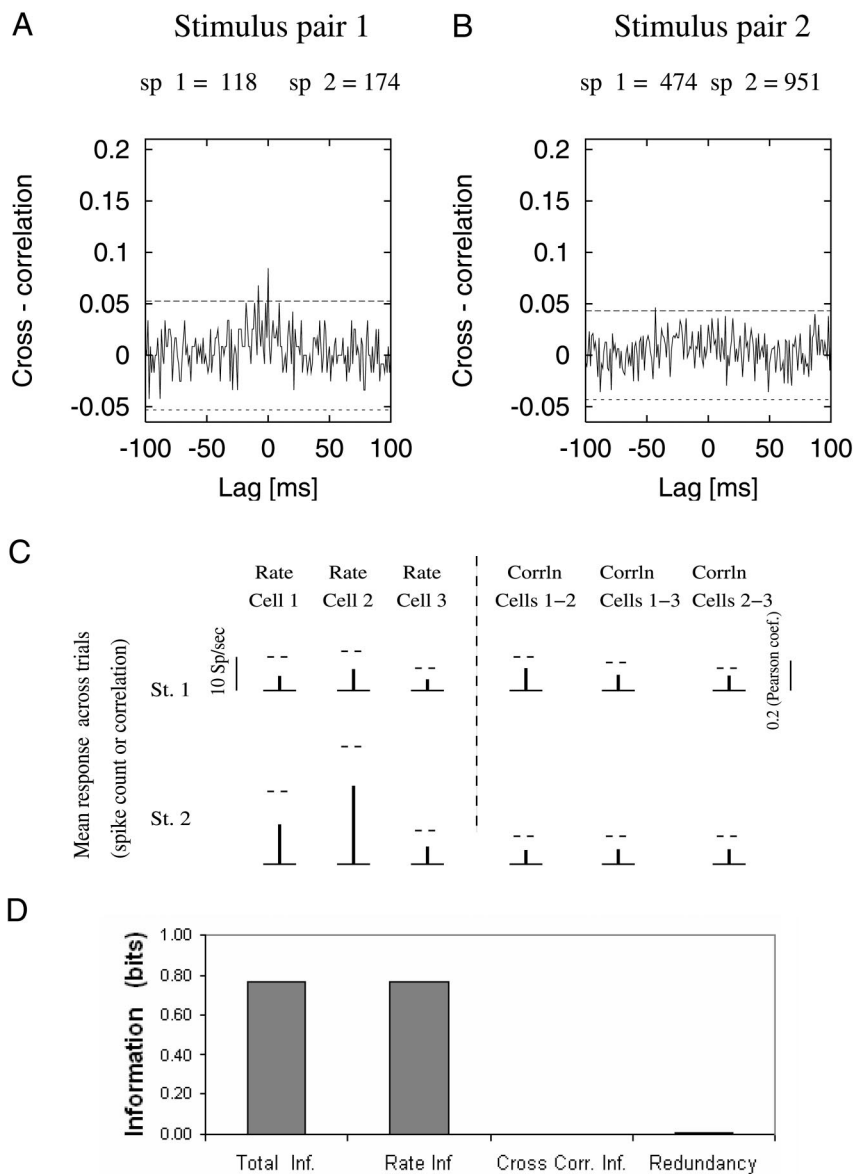
sp 1 = 474   sp 2 = 951

C

D

FIG. 4. Example of a set of neurons recorded with cross-correlations that are stimulus-dependent in investigation 1 in a plain background. Conventions as in Fig. 2 (experiment bj293b).

the first two columns does include the rate covariation redundancy that arises from the interaction between the within trial covariation of the response rates of the neurons and the correlations of their response tuning profiles (see Rolls et al. 2003b, 2004). Fourth, a further contribution to the generally lower information in the complex than the plain background (compare Tables 2 and 1) was the greater variability of the neuronal response in the complex background than in the plain background. To quantify this, we calculated the Fano factors (defined as the variance/mean rate, although calculated here from the slope of the variance with respect to the mean for all the cells to enable the especially variable high rates in the neuronal data to be taken into account, as they are relevant to the information calculation), and found an average in the plain background of 1.56 ± 0.04 (SE) and in the complex background of 1.72 ± 0.05 (in a 400-ms epoch; $P < 0.001$). (The corresponding Fano factors for the cross-correlation measure were 0.97 ± 0.07 and 1.02 ± 0.08.) The fact that there

was more variability (and less information) in the complex background is attributable to the low discriminability of the objects against the complex background (see Fig. 1). Indeed, there was behavioral evidence for the latter, in that the mean latency for the first correct touch of a stimulus was 615 ms with a blank background and 784 ms in the complex background ($P < 0.02$, t-test).

Table 2 and Fig. 6 (*bottom*) summarize the results for the same series of experiments, but with an analysis epoch of 100 ms. The same conclusions as those evident from Table 1 with the 400-ms epoch can be made. In addition, comparison of Table 2 with Table 1 shows that, on average, in a 100-ms epoch, 37.0% of the rate information relative to that in a 400-ms epoch was obtained (in a plain background). The corresponding figure for the complex background is 36.4%. Thus much of the information is available in quite short analysis periods.

Complex background

A Stimulus pair 1

sp 1 = 141   sp 2 = 325

B Stimulus pair 2
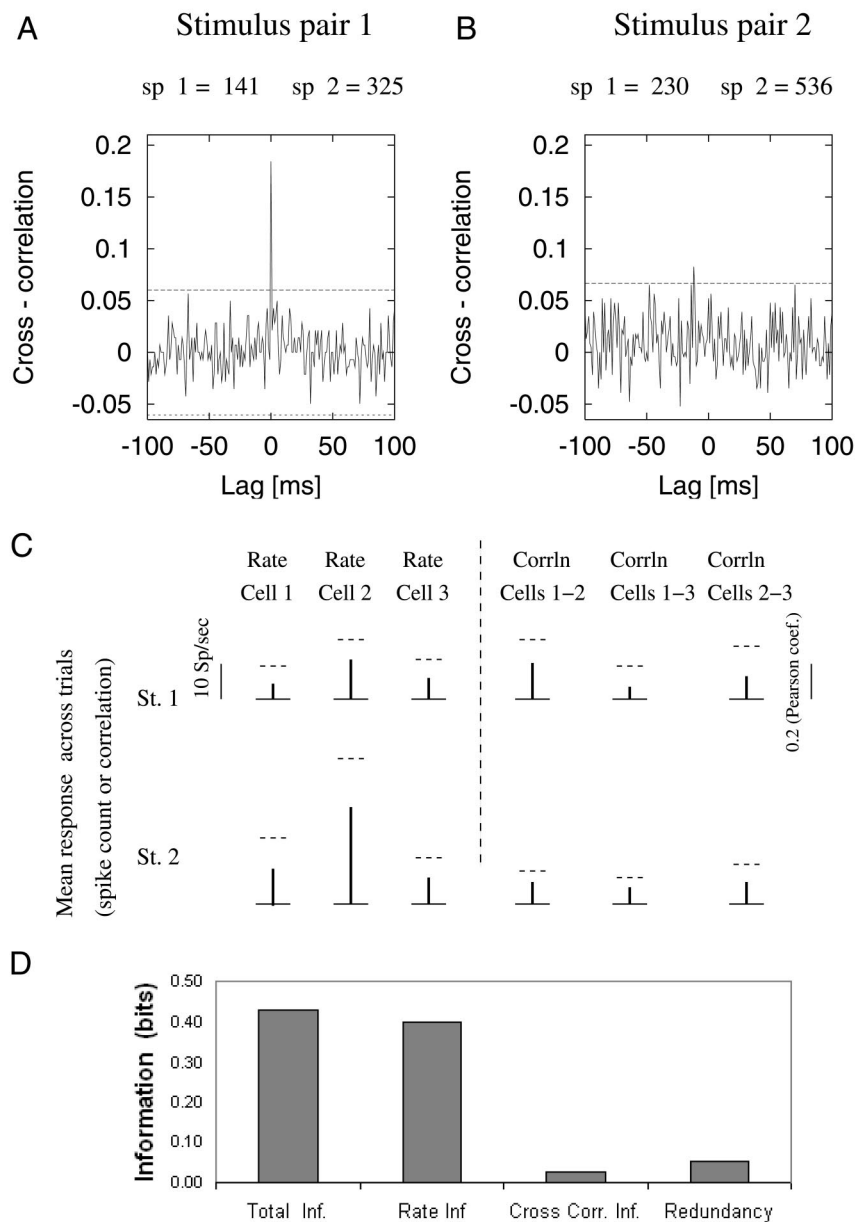
sp 1 = 230   sp 2 = 536

C

D

FIG. 5. Example of a set of neurons recorded with cross-correlations that are stimulus-dependent in investigation 1 in a complex background. Conventions as in Fig. 2. *A*: peak in the cross-correlogram located at lag 0 is clearly present for the case of 1 of the pairs of stimuli (*1*). There is no cross-correlation at this lag for stimulus pair 2. No clear cross-correlations were present for the same neuron with the plain background (see Fig. 4) (experiment bj293b).

We did apply the method used in investigation 2 of measuring the information about which image was being viewed in a simultaneously presented pair of images by taking data epochs only when the monkey was looking at one or the other of the images. Very little information was available about which of a particular pair of images was being viewed, consistent with the fact that, in investigation 1, at least some of the simultaneously recorded neurons were preselected to have similar firing rates to each member of a simultaneously presented pair. This finding is consistent with the overall result found for inferior temporal cortex neurons based on the results of both investigation 1 and investigation 2, that the information is available mainly in the rates and much less in any SDS that may be present. Indeed, we show in investigation 2 that there is information about which of two members of a simultaneously presented pair of images in a complex natural scene is being viewed, provided that there are firing rate differences to the two images.

Investigation 2

It was possible to complete 30 experiments in which with 2–4 electrodes, 2–5 neurons were simultaneously recorded for >80 trials while the monkey performed the visual discrimination task, touching the screen on every trial to obtain rewards if the correct image of the two being shown on the screen was touched. The total number of neurons in the sample was 89. In investigation 2, there was one pair of stimuli, and one stimulus was selected to be effective for one or more of the neurons, and the other stimulus was selected to be ineffective for one or more of the neurons. In different experiments, either the effective stimulus, or the ineffective stimulus, was rewarded.

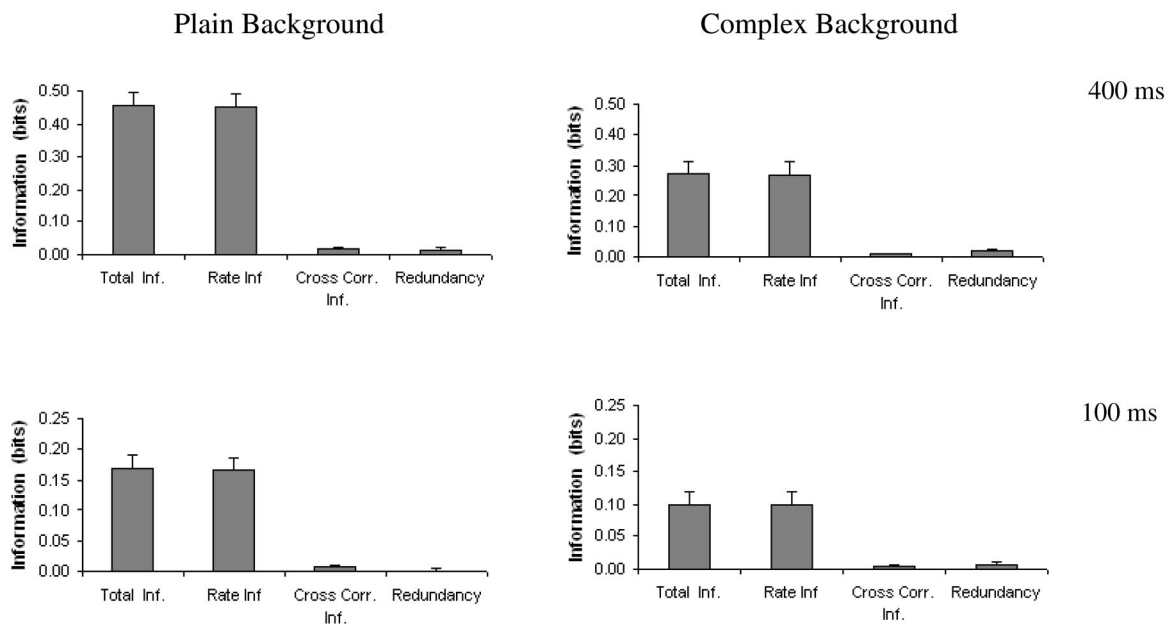## Plain Background
Complex Background

FIG. 6. Investigation 1. Average across 31 experiments each with 2–4 simultaneously recorded neurons of different components of the information about which stimulus was shown. Separate summaries are shown for 400- and 100-ms epochs.

Part of the interest of investigation 2 was that, in the image being shown, only one of the objects was effective for one (or more) of the neurons, and so it was possible to address how looking at one versus the other in a scene provided information that discriminated between the two objects in a single scene.

An example of the data acquisition with this experimental design is shown in Fig. 7. Eye positions and neuronal response data collection during the performance of the visual search task for two simultaneously recorded neurons are shown. Separate traces show the distance of the eyes from the target (rewarded) search object ($S+$) and from the distractor object ($S-$). Rastergrams for two simultaneously recorded neurons are shown above, with each vertical line representing an action potential from a neuron. The visual display was switched on at *time 0*. It can be seen that the neuron labeled 21 responded while the monkey looked at the $S+$ and fired less when the monkey looked at the $S-$. Conversely, neuron 31 fired rapidly while the monkey looked at the $S-$ and fired less when the monkey looked at the $S+$. There was less firing of this neuron when the monkey was fixating the other stimulus (which in this case was the $S-$). The neuronal activity in 100-ms epochs was collected for each of the stimuli while the monkey was looking with the eyes still within 3° of the center of each stimulus. [There could be several such epochs on a single behavioral trial. The epoch of data collection was delayed by 100 ms from the relevant eye position values to allow for the fact that inferior temporal cortex neurons have response latencies in the order of 100 ms (Baylis et al. 1987) and respond in a complex background ~100 ms after the eyes land on an effective target, as shown in Fig. 7 and by Rolls et al. 2003a.]

Table 3 and Fig. 8 summarize the data across all experiments in investigation 2 using the 100-ms analysis epoch, shown separately for plain and complex backgrounds. First, it is clear that, on average, across the 30 experiments, the information related to the firing rate (0.056 bits) was greater than the stimulus-dependent cross-correlation information (0.008 bits; for the plain background). Expressed at a percentage of the total information (0.061 bits), the rate information thus provided 91.3%. The SDS-related information provided 13.6% of the total information, although only 8.7% was independent of the firing rate–dependent information. This difference is also evident in the complex background (average rate information across experiments = 0.039 bits and average stimulus-dependent cross-correlation information = 0.005 bits). In the complex background, expressed as a percentage of the total information (0.041 bits), the rate information provided 94.4%. The SDS-related information provided 11.3% of the total information, although only 5.6% was independent of the firing rate–dependent information. [The rate and total information were lower in investigation 2 than investigation 1 (cf. Tables 2 and 3), and perhaps this was not surprising, because the information being measured in investigation 2 between two stimuli simultaneously presented sufficiently close for the receptive fields to overlap. Indeed, in the complex scene in investigation 2, the mean firing rate to the more effective stimulus of a pair was 26.2 spikes/s and to the less effective was 20.7 spikes/s compared with 27.4 and 16.3 spikes/s, respectively, for investigation 1.] Second, there was somewhat less information in the complex background. [We did not measure the rate covariation redundancy in investigation 2 because we used the maximum likelihood decoding method, because this has the advantage of high sensitivity when information values are low; they were in investigation 2 partly because we used a short analysis epoch, partly because some of the 100-ms epochs were after the neuron had already been firing for >100 ms to the stimulus when the firing rates tend to be a little lower, as shown in typical peristimulus time histograms (Rolls and Deco 2002), and partly because the objects had low discriminability against the complex background. The full probability estimation decoding method used in investigation 1 uses the full stimulus-response probability table, and in using more of the values, provides a more smoothed estimate of the information that we have found to be useful when quantifying redundancy (Franco et al. 2004).]
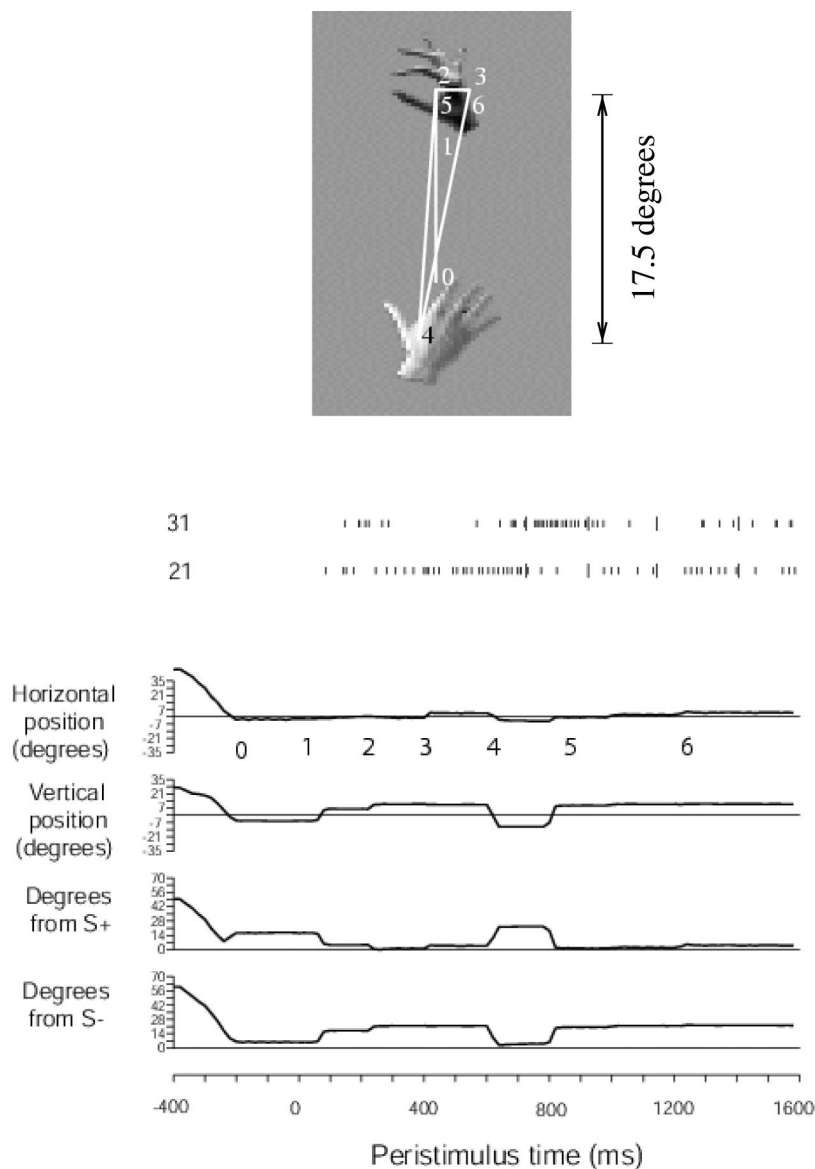
FIG. 7. Investigation 2. Eye positions and neuronal response data collection during the performance of visual search task for 2 simultaneously recorded neurons. Horizontal and vertical eye position traces are calibrated with respect to the center of the screen in degrees (with −35° horizontal and −35° vertical being the lower left of the screen). Separate traces show the distance of the eyes from the target (rewarded) search object (S+) and from the distractor object (S−). Rastergrams for the 2 simultaneously recorded neurons are shown above, with each vertical line representing an action potential from a neuron. Visual display was switched on at *time 0*. Neuron 21 responded when the monkey looked at the S+, and neuron 31 when the monkey looked at the S−. The monkey made multiple touches of the S+ (each indicated by a long vertical line) of the object to obtain fruit juice. Diagram at *top* shows a reconstruction of eye positions on an enlarged part of display with numbers keying each position on the reconstruction to eye position plots shown below.

The sets of simultaneously recorded neurons were not especially selected to have rate versus synchrony-related information. [During each experiment, it was known that 1 to 2 of the typically 5 neurons did have different firing rates to the stimuli (which is quite typical for inferior temporal cortex neurons), and they, and the other neurons, could have had stimulus-dependent correlations (although this was not evaluated at the time of the recording) and were therefore included in the information theoretic analyses.] If we ask the biased question of, for just those experiments in which the neurons conveyed any stimulus-dependent information, what proportion was stimulus-dependent and what was rate-dependent, the results are not strikingly different from the proportions above. For example, in 16 experiments with the complex scene from Table 3 in which the neuron ensemble had any SDS-related information, expressed at a percentage of the total information (0.071 bits), the rate information thus provided 86%. The SDS-related information provided 22% of the total information, although only 14% was independent of the firing rate–dependent information. We further note that the information (both rate and SDS) was present to similar extents early on and later in a trial, so that possible decision-related factors occurring early in a trial (in the 1st fixation on an image) compared with later in the trial did not have clear effects on the responses of the inferior temporal cortex neurons described here.

Although the results presented so far were in one monkey, and thus the results from different experiments can be directly compared, we have been able to establish that the results are replicable, in that, in a second monkey, we have been able to perform seven further experiments in which 23 further neurons were analyzed in investigation 2. The results were very similar to those reported above. In particular, for the plain background, the rate information provided 92.7% of the total information. The SDS-related information provided 15.2% of the total information, although only 7.3% was independent of the firing rate–dependent information. In the complex background, the rate information provided 98.0% of the total information. The SDS-related information provided only 4.6% of the total information, and only 2.0% was independent of the firing rate information. Thus the findings on the relative contributions of

TABLE 3. *Experiment 2: the contributions in bits of different components of the information extracted using a decoding algorithm*

	Rate Info	Stim-dep. Correlation Info	Total Info	Rate Info	Stim-dep. Correlation Info	Total Info
	Plain	Plain	Plain	Complex	Complex	Complex
bj215	0.04	0	0.04	0.08	0.01	0.08
bj244b	0	0.03	0	0	0	0
bj247a	0.04	0.01	0.05	0	0	0
bj253f	0.24	0	0.24	0.04	0	0.04
bj253g	0.16	0	0.16	0.07	0	0.07
bj253i	0.14	0.01	0.14	0.03	0.03	0.03
bj253j	0.36	0.02	0.36	0.19	0.02	0.19
bj270a	0.03	0	0.03	0.1	0	0.1
bj270b	0	0	0	0.06	0	0.06
bj272	0.03	0	0.03	0	0	0
bj273a	0.06	0	0.06	0.05	0	0.05
bj273b	0.09	0	0.09	0	0	0
bj273c	0.08	0.01	0.08	0.03	0	0.03
bj274b	0.01	0	0.01	0.08	0.01	0.08
bj275a	0.1	0.08	0.18	0	0	0
bj276	0	0	0	0.05	0	0.05
bj277	0.06	0.01	0.07	0.09	0	0.09
bj280a1	0.02	0.02	0.02	0.06	0	0.06
bj280b	0	0.03	0.03	0.02	0	0.02
bj284	0.04	0	0.04	0	0.03	0.03
bj285a	0	0	0	0.05	0	0.05
bj285b	0	0	0	0	0.03	0.03
bj288	0	0.01	0.01	0	0.01	0.01
bj289a	0.08	0.01	0.09	0	0	0
bj290	0.03	0	0.03	0.05	0	0.05
bj291	0	0	0	0.03	0	0.03
bj292	0.01	0.01	0.02	0.02	0	0.02
bj320	0.03	0	0.03	0.05	0	0.05
bj365a	0.02	0	0.02	0.02	0	0.02
bj387a	0.01	0	0.01	0	0	0
Averages	0.056	0.008	0.061	0.039	0.005	0.041
Percentage of total	91.3%	13.6%	100%	94.4%	11.3%	100%

Maximum likelihood, with a criterion of $I > 0.01$ bits. All the information values shown are for 100-ms epochs.

firing rate and SDS effects to the total information have been confirmed in two different monkeys, in which 221 neurons were analyzed in 68 different experiments.

Figure 9 shows the recording sites. Reconstructed histological coronal sections show, with filled circles, the sites at which the neurons analyzed in this paper were recorded. Numbers below the sections indicate the distance (in mm) posterior to the sphenoid bone reference point (which is at approximately the anterior-posterior level of the anterior commissure), and these distances are further shown in the *top left* of the figure in the lateral view. A full coronal section is shown at the *top right*, and the area of cortex investigated in this study is indicated by the shaded region encompassing the STS and the lateral portion of the inferior temporal gyrus (IT). Recording tracks were made over an extensive portion of the inferior temporal cortex, from the upper and lower banks and fundus of the superior temporal sulcus, through the middle temporal gyrus to just lateral to the middle temporal sulcus. As can be seen in Fig. 9, the cells are distributed from lateral of the middle temporal sulcus to the lower bank of the STS, and the investigated area of cortex is indicated by the shaded bounding box in the coronal section shown in the *top right* of Fig. 9.

DISCUSSION

Both investigations 1 and 2 used displays with two objects in complex scenes, in which the monkey had to segment the two objects from the background, and from each other, and after object recognition, reach to touch the object that was associated with reward. This is an interesting situation in which to test neural encoding because the visual system is operating under natural visual conditions in which, to the extent that this is required with natural vision, features must be segmented from the background and bound together if the features are part of the same object. Under these natural visual conditions, we found that inferior temporal cortex neurons encode objects in such a way that 93–99% of the total information is available from the firing rate (or spike counts), with the stimulus-dependent cross-correlations adding only 1–6% of independent information. Even when only the information available from the stimulus-dependent cross-correlations was measured, it amounted to only 3–11% of the total information. Thus spike counts across a population of neurons provide much more information than stimulus-dependent cross-correlations, and to the extent that stimulus-dependent cross-correlations provide information, it is not fully independent from the rate information. We note that if more neurons are recorded simultaneously, that the numbers of pairwise cross-correlations increases according to $n(n - 1)/2$ where n is the number of neurons, and that although there is some potential for the information available from stimulus-dependent cross-correlations to rise, there is, in fact, the likelihood that all these correlations would not be independent of each other and thus

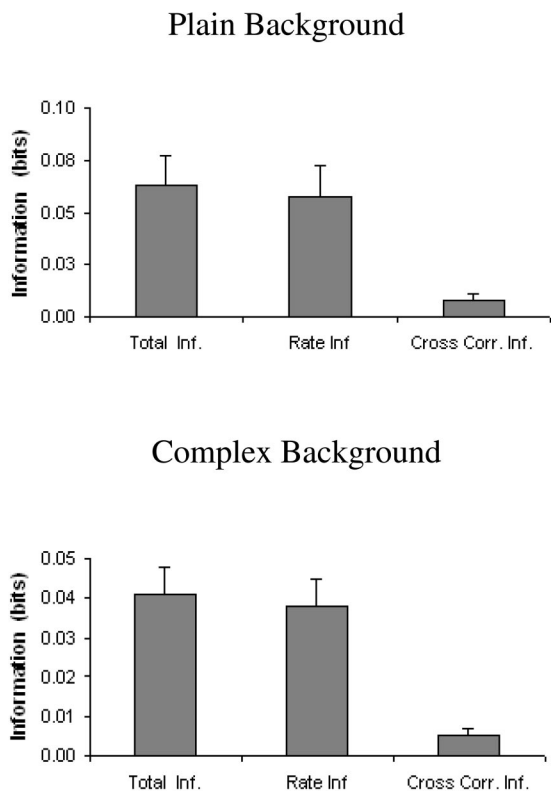## Plain Background

Complex Background

FIG. 8. Investigation 2. Average across 30 experiments each with 2–4 simultaneously recorded neurons of different components of the information about which stimulus was shown in 100-ms epochs in which the monkey was fixating 1 of the 2 stimuli.

might not lead to a rapid increase in the information available from the stimulus-dependent cross-correlations. In contrast, we do expect that the information from the firing rates will grow approximately linearly with the number of neurons considered (Abbott et al. 1996; Rolls et al. 2004).

The information theoretic method we used for measuring the relative contributions of spike counts and stimulus-dependent neuronal synchrony in populations of neurons shows how these contributions can be quantitatively compared (Franco et al. 2004; for earlier approaches, see Gawne and Richmond 1993; Hatsopoulos et al. 1998; Oram et al. 1998; Reich et al. 2001; Rolls et al. 2003b, 2004). In previous studies, it has been shown that SDS is a property of neuronal firing under particular conditions (e.g., Singer 1999, 2000); however, this is not sufficient to show how quantitatively important it is. To answer that question, it is necessary to know how much information can be gained on a single trial from SDS, as the variability in the SDS is extremely relevant to how much information can be gained from it. An important conclusion from the findings reported in this paper is that, even when the SDS may look strong in a cross-correlogram (and even may look smooth if hundreds of trials are used), there may be rather little information available from SDS on a single trial basis. In contrast, much more information is available on a single trial basis from the spike counts. Even if some earlier stage of the visual system than the inferior temporal visual cortex might perform feature binding by SDS, we note that this general point, about how much can be learned from spike counts versus SDS even when the latter is present, is very important. However, the inferior temporal cortex, with its receptive fields that are large

enough to encompass whole objects and where neurons can respond to features of objects as well as to objects (Desimone et al. 1984; Gross et al. 1972; Perrett et al. 1982, 1992; Rolls et al. 1994; Vogels 1999), does seem a candidate for a visual cortical area in which feature binding is needed and where the SDS hypothesis can be tested.

Although we have discussed the finding so far when objects are being segmented from natural backgrounds as well as separated from each other, we note that, in the plain background, most of the total information came from the spike counts (98%), leaving only ~2% of independent information from SDS.

The information theoretic approach we used also allowed us to show that there was little rate covariation redundancy between the information provided by the spike counts of the simultaneously recorded neurons, making spike counts a powerful population code. In the complex background, the redundancy averaged 6.3%, and in the plain background, 1.2% (see Tables 1 and 2). There was less information, on average, from the groups of neurons about the stimuli in the complex than in the plain background. This probably arose from lower single cell information in the complex background, evident as smaller firing rate differences between effective and less effective stimuli when tested in the complex versus the plain background. This is probably related to the fact that the stimuli we used in this experiment were intentionally not highly discriminable from the complex background (see Fig. 1), to make the objects difficult to segment, to increase the chance that SDS, if used in segmentation, would be measured in this investigation. The greater rate covariation redundancy with the complex background probably was related to the greater similarity of the responses of the neuronal populations to the two stimuli in the
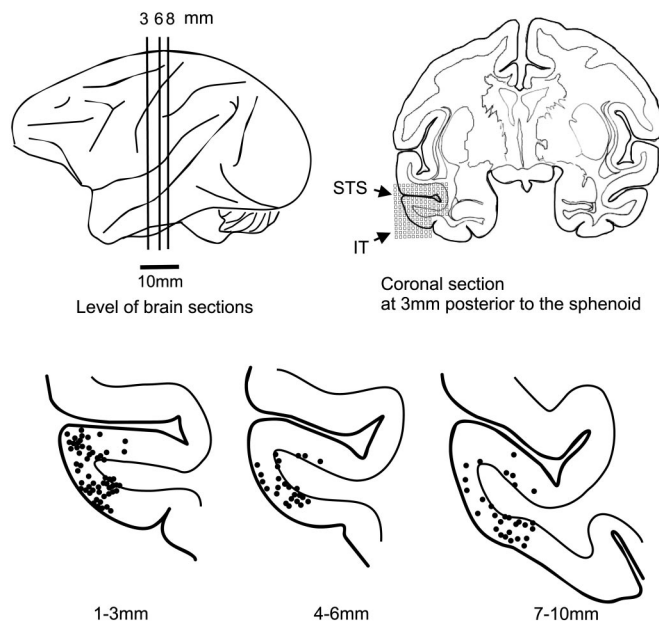

FIG. 9. Reconstructed histological coronal sections show the neuronal recording sites (filled circles). Numbers below sections indicate distance (in mm) posterior to the sphenoid bone reference point (which is at approximately the anterior-posterior level of the anterior commissure), and these distances are further shown in the *top left* of the figure in the lateral view. A full coronal section is shown at the *top right*, and the area of cortex investigated in this study is indicated by the shaded region encompassing the superior temporal sulcus (STS) and the lateral portion of the inferior temporal gyrus (IT).

complex background, due to the smaller firing rate response differences in the complex background (i.e., the profiles of the responses of the population of neurons to the two stimuli become more similar in the complex background).

Comparison of Table 2 with Table 1 shows that, on average, in a 100-ms epoch, 37.0% of the rate information relative to that in a 400-ms epoch was obtained (in a plain background). The corresponding figure for the complex background is 36.4%. For the case of SDS-related information, the values are 44% for the plain background and 56% for the complex background. Thus much of the information is available in quite short analysis periods. The extension to earlier work (Tovee and Rolls 1995) is that this is now supported by the new recordings and analyses with simultaneously recorded populations of neurons.

It is interesting and important that, when visual search is being performed for objects shown in complex scenes, the tuning of inferior temporal cortex neurons to the objects remains relatively unaffected (Rolls et al. 2003a; Sheinberg and Logothetis 2001), although the receptive fields become much smaller in a complex scene than against a plain background (Rolls et al. 2003a). These results are complemented by the findings of DiCarlo and Maunsell (2000) and Missall et al. (1999) that inferotemporal cortex neurons respond similarly to an effective shape stimulus for a cell even if some distractor stimuli are present a few degrees away. This finding might be called "background invariance", to capture the point that the tuning of many inferior temporal cortex neurons is invariant when the stimuli are shown against a background. This study quantifies for the first time the effects on the amount of information represented in plain versus complex scenes. The total information available is somewhat less in complex (0.041 bits in a 100-ms epoch in investigation 2 as shown in Table 3) than in plain backgrounds (0.061 bits). This was true to a similar extent for the rate and the synchrony-related information. With respect to the rate information, there was a small reduction of firing rate to the effective stimulus in the complex scene (28.5-25.4 spikes/s), but some of the reduction of information must have been related to increased trial by trial variability.

One of the conclusions of this paper is that little stimulus-dependent information from the cross-correlations was available about which stimulus was shown from the neurons recorded in the inferior temporal visual cortex, even under natural vision conditions. Could this be because the code is so sparse that it is difficult to detect and might require simultaneous recordings from very large numbers of neurons to be detected? Although this is certainly possible, we would argue that considerable information was available from the spike counts of the simultaneously recorded neurons about which stimulus was shown, and that this information could be easily decoded by receiving neurons, which might be more difficult if the code was very sparse. It certainly remains a possibility that SDS, perhaps measured with different techniques, and in perhaps other more artificial visual testing conditions, might be important in information encoding. We have measured synchrony with the normal cross-correlation method, and under natural vision conditions, and this is why the results are of interest. The other main conclusion is that considerable information is available on the spike counts (or firing rates) of inferior temporal cortex neurons about the object being viewed in a complex scene and that there is little "rate covariation" redundancy across at least small numbers of simultaneously recorded neurons. Furthermore, we note that the rate information we described in this paper is from the number of spikes available from each of a large number of neurons, such as might be presented to a receiving neuron, in a short epoch. In this paper, we have analyzed an epoch as short as 100 ms, and the results are likely to generalize to shorter time intervals such as 20 ms, given what we know about the encoding of information by single neurons (Tovee and Rolls 1995). Thus we do not envision that a receiving neuron would need to make an accurate measurement over, for example, 500 ms of the firing rates of its inputs. Instead, many sending neurons would each provide zero, one, or two spikes in a 20-ms period, a typical integration period for a receiving neuron, and this would be how the firing rate information that we have shown is available is being used.

REFERENCES

Abbott LF, Rolls ET, and Tovee MJ. Representational capacity of face coding in monkeys. *Cereb Cortex* 6: 498–505, 1996.

Baylis GC, Rolls ET, and Leonard CM. Functional subdivisions of temporal lobe neocortex. *J Neurosci* 7: 330–342, 1987.

Booth MCA and Rolls ET. View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb Cortex* 8: 510–523, 1998.

Cover TM and Thomas JA. *Elements of Information Theory.* New York: Wiley, 1991.

Dan Y, Alonso JM, Usrey WM, and Reid RC. Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature Neurosci* 1: 501–507, 1998.

Desimone R, Albright TD, Gross CG, and Bruce C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4: 2051–2062, 1984.

DiCarlo JJ and Maunsell JHR. Form representation in monkey inferotemporal cortex is virtually unaltered by free viewing. *Nature Neurosci* 3: 814–821, 2000.

Feigenbaum JD and Rolls ET. Allocentric and egocentric spatial information processing in the hippocampal formation of the behaving primate. *Psychobiology* 19: 21–40, 1991.

Franco L, Rolls ET, Aggelopoulos NC, and Treves A. The use of decoding to analyze the contribution to the information of the correlations between the firing of simultaneously recorded neurons. *Exp Brain Res* 155: 370–384, 2004.

Gawne TJ and Richmond BJ. How independent are the messages carried by adjacent inferior temporal cortical neurons? *J Neurosci* 13: 2758–2771, 1993.

Gross CG, Rocha Miranda CE, and Bender DB. Visual properties of neurons in inferotemporal cortex of the macaque. *J Neurophysiol* 35: 96–111, 1972.

Hatsopoulos NG, Ojakangas CL, Paninski L, and Donoghue JP. Information about movement direction obtained by synchronous activity of motor cortical neurons. *Proc Natl Acad Sci USA* 95: 15706–15711, 1998.

Judge SJ, Richmond BJ, and Chu FC. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20: 535–538, 1980.

Kayser C, Salazar RF, and Konig P. Responses to natural scenes in cat V1. *J Neurophysiol* 90: 1910–1920, 2003.

Malsburg CVD. A neural architecture for the representation of scenes. In: *Brain Organization and Memory: Cells, Systems and Circuits,* edited by McGaugh JL, Weinberger NM, and Lynch G. New York: Oxford University Press, 1990, p. 356–372.

Missall M, Vogels R, Chao-Yi L, and Orban GA. Shape interactions in inferior temporal neurons. *J Neurophysiol* 82: 131–142, 1999.

Oram MW, Foldiak P, Perrett DI, and Sengpiel F. The 'ideal homunculus': decoding neural population signals. *Trends Neurosci* 21: 259–265, 1998.

Oram MW, Hatsopoulos NG, Richmond BJ, and Donoghue JP. Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J Neurophysiol* 86: 1700–1716, 2001.

Panzeri S, Schultz SR, Treves A, and Rolls ET. Correlations and the encoding of information in the nervous system. *Proc Roy Soc B Lond* 266: 1001–1012, 1999.

Panzeri S and Treves A. Analytical estimates of limited sampling biases in different information measures. *Network* 7: 87–107, 1996.

Perrett DI, Hietanen JK, Oram MW, and Benson PJ. Organisation and functions of cells responsive to faces in the temporal cortex. *Philo Trans Roy Soc Lond* 335: 23–30, 1992.

Perrett DI, Rolls ET, and Caan W. Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47: 329–342, 1982.

Reich DS, Mechler F, and Victor JD. Independent and redundant information in nearby cortical neurons. *Science* 294: 2566–2568, 2001.

Rolls ET. Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron* 27: 205–218, 2000.

Rolls ET, Aggelopoulos NC, Franco L, and Treves A. Information encoding in the inferior temporal cortex: contributions of the firing rates and correlations between the firing of neurons. *Biol Cybern* 90: 19–32, 2004.

Rolls ET, Aggelopoulos NC, and Zheng F. The receptive fields of inferior temporal cortex neurons in natural scenes. *J Neurosci* 23: 339–348, 2003a.

Rolls ET and Deco G. *Computational Neuroscience of Vision.* Oxford: Oxford, 2002.

Rolls ET, Franco L, Aggelopoulos NC, and Reece S. An information theoretic approach to the contributions of the firing rates and the correlations between the firing of neurons, *J Neurophysiol* 89: 2810–2822, 2003b.

Rolls ET, Judge SJ, and Sanghera M. Activity of neurones in the inferotemporal cortex of the alert monkey. *Brain Res* 130: 229–238, 1977.

Rolls ET, Sanghera MK, 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 ET and Tovee MJ. Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J Neurophysiol* 73: 713–726, 1995.

Rolls ET, Tovee MJ, Purcell DG, Stewart AL, and Azzopardi P. The responses of neurons in the temporal cortex of primates, and face identification and detection. *Exp Brain Res* 101: 473–484, 1994.

Rolls ET, Treves A, and Tovee MJ. The representational capacity of the distributed encoding of information provided by populations of neurons in the primate temporal visual cortex. *Exp Brain Res* 114: 149–162, 1997.

Shadlen M and Movshon J. Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron* 24: 67–77, 1999.

Shannon CE. A mathematical theory of communication. *AT&T Bell Lab Tech J* 27: 379–423, 1948.

Sheinberg DL and Logothetis NK. Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. *J Neurosci* 21: 1340–1350, 2001.

Singer W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24: 49–65, 1999.

Singer W. Response synchronisation: a universal coding strategy for the definition of relations. In: *The New Cognitive Neurosciences,* edited by Gazzaniga M. Cambridge, MA: MIT Press, 2000, p. 325–338.

Singer W and Gray CM. Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci* 18: 555–586, 1995.

Tamura H and Tanaka K. Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cereb Cortex* 11: 384–389, 2001.

Tanaka K. Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19: 109–139, 1996.

Tovee MJ and Rolls ET. Information encoding in short firing rate epochs by single neurons in the primate temporal visual cortex. *Vis Cogn* 2: 35–58, 1995.

Treves A. Information coding in higher sensory and memory areas. In: *Handbook of Biological Physics,* edited by Moss F and Gielen S. Amsterdam: Elsevier, 2000, p. 803–829.

Treves A and Panzeri S. The upward bias in measures of information derived from limited data samples. *Neural Comput* 7: 399–407, 1995.

Vogels R. Effect of image scrambling on inferior temporal cortical responses. *Neuroreport* 10: 1811–1816, 1999.

Journal of Neurophysiology