

Scene perception: inferior temporal cortex neurons encode the positions of different objects in the scene

Nikolaos C. Aggelopoulos and Edmund T. Rolls

University of Oxford, Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK

Keywords: multiple object perception, object recognition, rhesus monkey, scene encoding, spatial vision, translation invariance

Abstract

Inferior temporal cortex (IT) neurons have reduced receptive field sizes in complex natural scenes. This facilitates the read-out of information about individual objects from IT, but raises the question of whether more than the single object present at the fovea is represented by the firing of IT neurons, as would be important for whole scene perception in which several objects may be located without eye movements. Recordings from IT neurons with five simultaneously presented objects, each subtending 7°, with one object at the fovea and the other four centred 10° eccentrically in the parafovea, showed that although 38 IT neurons had their best response to an effective stimulus at the fovea, eight IT neurons had their best response to an object when it was located in one or more of the parafoveal positions. Moreover, of 54 neurons tested for asymmetric parafoveal receptive fields, 35 (65%) had significantly different responses for different parafoveal positions. The asymmetry was partly related to competition within the receptive fields, as only 21% of the neurons had significant asymmetries when tested with just one object present located at the same parafoveal positions. The findings thus show that some evidence is conveyed by a population of IT neurons about the relative positions of several simultaneously presented objects in a scene extending well into the parafovea during a single fixation, and this is likely to be important in whole scene perception with multiple objects, including specifying the relative positions of different objects in a scene.

Introduction

Inferior temporal cortex (IT) neurons of macaques have responses that provide information about objects or faces (e.g. see Gross *et al.*, 1972; Perrett *et al.*, 1982; Rolls & Baylis, 1986; Desimone, 1991; Rolls, 1992, 2000, 2005; Tanaka, 1996; Booth & Rolls, 1998; Vogels, 1999; Rolls & Deco, 2002). Much visual neurophysiology is conducted with one visual stimulus present in an otherwise blank visual scene (Hubel & Wiesel, 1962; Gross *et al.*, 1985), and even in most studies of selective attention, only two small visual stimuli are present (Chelazzi *et al.*, 1993; Desimone & Duncan, 1995; Chelazzi, 1998; Chelazzi & Corbetta, 2000). When one visual stimulus is present, IT neurons typically have large receptive fields of 50° or more (Gross *et al.*, 1972; Tovee *et al.*, 1994; Op De Beek & Vogels, 2000; Rolls *et al.*, 2003). However, a potential problem then is that IT neurons might not support object selection by subsequent brain areas, as different neurons tuned to different objects but all with large receptive fields might respond simultaneously, so that the IT output might appear as a ‘tower of Babel’. Therefore, it was an interesting discovery that in complex natural scenes IT receptive fields decreased to approximately 10° radius and were close to the fovea (Sheinberg & Logothetis, 2001; Rolls *et al.*, 2003). These much smaller receptive fields close to the fovea reduce the binding problem, as the output of the system is primarily about what is at or close to the fovea. Thus, brain regions such as the orbitofrontal cortex and amygdala, which receive outputs from IT, can utilize easily the information that they receive, as it is

primarily about the object at the fovea. There is also evidence that under conditions in which macaques have been trained to make conditional eye movement responses between multiple small (0.6°) geometric shapes arranged in a row in a visual fixation task, IT neurons respond differently to stimuli at the fovea or 1.5° contralaterally (DiCarlo & Maunsell, 2003). In addition, some interactions between stimuli in the receptive fields of IT neurons have been found (Sato, 1989, 1995; Miller *et al.*, 1993; Rolls & Tovee, 1995a; Missall *et al.*, 1999).

However, if no eye movement is allowed, do IT neurons as a population report on more than just a single object at the fovea? If they did, and the relative spatial position of different objects in the parafovea were encoded by the population, this would be very useful in helping to build whole scene representations, and would supplement the use of multiple saccades to different positions in the scene. Consistent with this possibility, humans can report several aspects of a scene even with an exposure (100 ms) which is too short to allow eye movements to be used (Biederman, 1972; Biederman *et al.*, 1974).

In the experiments described here, we investigated the response properties of macaque IT neurons when several objects were present simultaneously at foveal and parafoveal locations, to show whether IT neurons provide information about more than the stimulus at the fovea.

Methods

The experiment was designed to provide evidence on how IT neurons respond in complex scenes with multiple objects present.

Correspondence: Professor E. T. Rolls, as above.

E-mail: Edmund.Rolls@psy.ox.ac.uk

Received 6 July 2005, revised 4 October 2005, accepted 5 October 2005

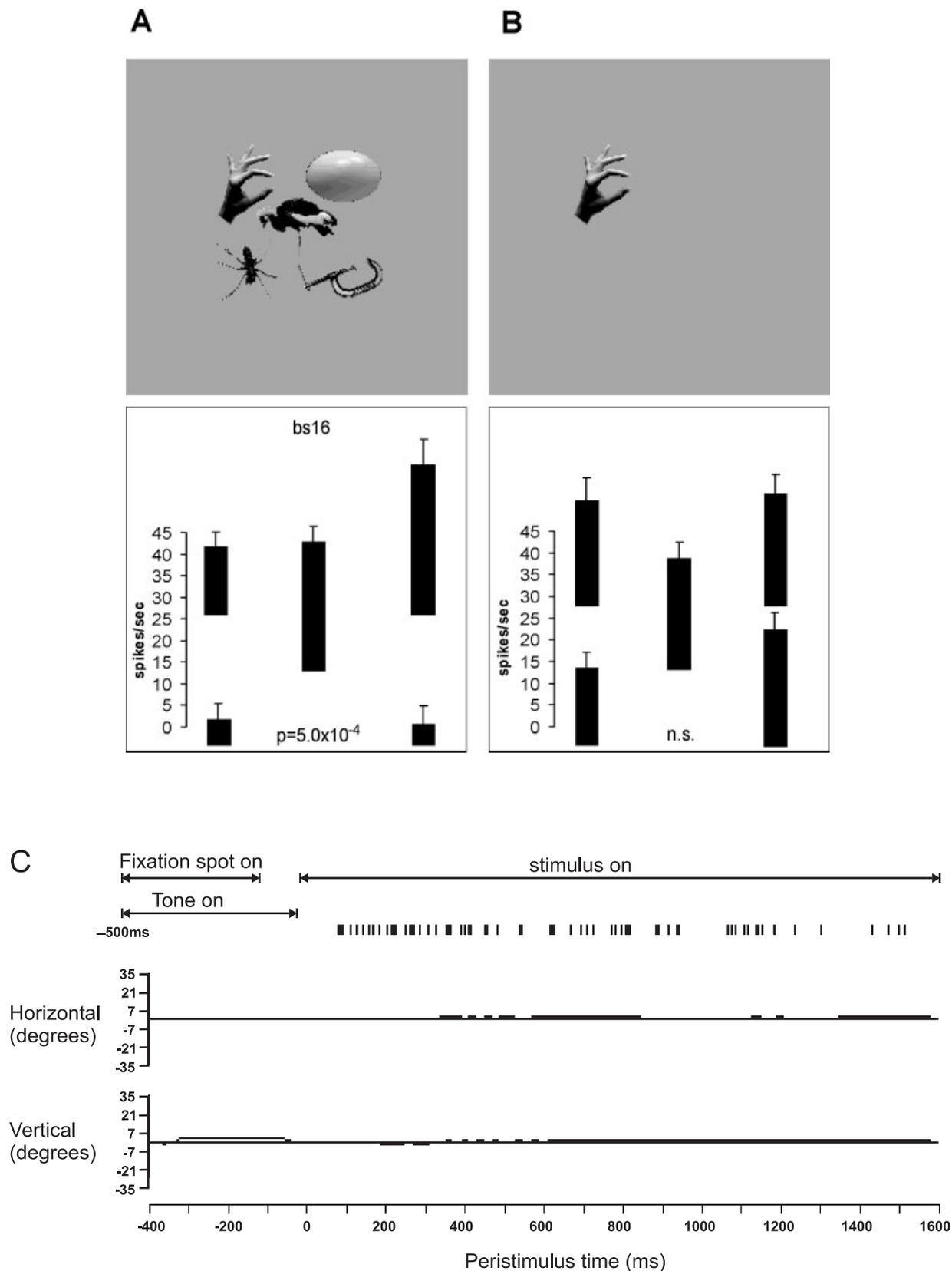


FIG. 1. (A) The responses (firing rate with the spontaneous rate subtracted, means \pm SEM) of one neuron when tested with five stimuli simultaneously present in the close (10°) configuration with the parafoveal stimuli located 10° from the fovea. (B) The responses of the same neuron when only the effective stimulus was presented in each position. The firing rate for each position is that when the effective stimulus for the neuron was in that position. The P -value is that from the ANOVA calculated over the four parafoveal positions. (C) An eye position recording to show that fixation at the centre of the screen was maintained throughout the trial. The visual stimulus appeared at time 0.

Subjects

The subjects were two macaque monkeys (*Macaca mulatta*, weight ~7–9 kg). All procedures, including preparative and subsequent ones, were carried out in accordance with the NIH Guide for the Care and Use of Laboratory Animals, and were licensed under the UK Animals (Scientific Procedures) Act 1986.

Overall design of the investigation

The experiment was designed to provide evidence on how IT neurons respond in complex scenes with multiple objects present, and used five stimuli simultaneously presented with one at the fovea and four in the surrounding parafovea so that quantitative and controlled data could be obtained. The stimuli were arranged as shown in Fig. 1, and the eyes were held still on a central object with a visual fixation task. This task ensured that the fovea was directed at the centre of the screen where the fixation spot was located at the start of each trial (see below). [For the purposes of this paper, the fovea is defined as a region approximately 2–5° in diameter that contains densely packed cones, and which is surrounded by the parafovea where cone density though still high does decrease (Polyak, 1941; Adler, 1959; Rolls & Cowey, 1970; Palmer, 1999).] This design had the advantage over a complex natural scene of enabling defined stimuli to be presented simultaneously in defined positions in a well-controlled way, and for the response of a neuron to be measured to each stimulus when presented alone. Moreover, the stimuli were chosen such that only one of the stimuli activated the particular neuron from which recordings were being made (as shown by ANOVAs and *post-hoc* analyses). This enabled the response of a neuron to be measured to its effective stimulus when it was present in each of the five positions, with non-effective stimuli for that particular neuron being shown at the other four locations. The stimuli were presented with randomly permuted positions on each trial, so that it was possible to measure the firing rate when the effective stimulus was at the fovea and at each of the four parafoveal positions. This enabled measurement of whether the neuron responded to the object only when it was at the fovea, or whether it also responded when the effective stimulus was at a parafoveal position and there was an ineffective stimulus for the neuron at the fovea. If there was a response in a parafoveal position, it was possible with the design to show whether the receptive field was symmetric or asymmetric with respect to the fovea, and how large the response was when the effective stimulus was in a parafoveal vs. a foveal position. One comparison condition included the same task but with only the effective stimulus in the display on any one trial, at one of the five positions. This provided evidence on whether the same neuron had a large receptive field when only one object was present in the scene. A second comparison condition ('far') included the five simultaneously presented stimuli, but now spaced out further from the fovea, to investigate how far the effects discovered extend from the fovea. The far condition, with the stimuli centred at 16° eccentricity in the parafoveal positions, was designed to allow a gap between the objects (see Fig. 3), to show how the receptive field interactions decrease as the stimuli are separated. The five simultaneously presented objects could be thought of as forming a complex scene, but at the same time as allowing the neuronal responses to the defined objects to be clearly measured when they were in different positions with respect to the fovea and to other stimuli.

Experimental protocol, and the stimuli

Images of five objects were presented simultaneously with the arrangement shown in Fig. 1. The objects were centred in a region

that subtended $7 \times 6^\circ$ at the retina (with a pixel resolution for each object of 64×64 and a greyscale resolution of 255), and the centre of the parafoveal objects was offset 10° diagonally from the fovea. [The spatial resolution of the objects was much greater than the typical spatial frequency requirement of IT cells, which respond well to stimuli provided that they contain frequencies up to approximately 8 cycles per stimulus or 16 pixels per stimulus (Rolls *et al.*, 1985), and was much more than sufficient to provide for very selective responses of the neurons to the different stimuli as shown by the ANOVAs.] In this way the parafoveal objects' edges were close to the edge of an object simultaneously presented at the fovea. This condition is termed the 'close' condition, because the objects were closely spaced. The stimuli were chosen in prior screening of a large set of stimuli of the type known to activate different inferior temporal cortex (Rolls & Tovee, 1995b), with examples illustrated in Fig. 1, such that only one of the stimuli activated the particular neuron from which recordings were being made. Activation was defined as a firing rate response that was more than 10 spikes/s above the spontaneous rate. An ineffective stimulus was defined as one which produced no significant change in firing as shown by the peristimulus time histogram. Many stimuli are ineffective for IT neurons, given the approximately exponential firing rate distribution of these neurons to a set of stimuli (Rolls & Tovee, 1995b; Baddeley *et al.*, 1997; Treves *et al.*, 1999). The refresh rate of the LCD computer monitor on which the stimuli were shown was 100 Hz.

The effective stimulus was presented at one of the five randomly permuted positions on each trial, so that it was possible to measure the firing rate when the effective stimulus was at the fovea and at each of the four parafoveal positions. The other stimuli were presented at random positions on each trial. This enabled measurement of whether the neuron responded to the object only when it was at the fovea, or whether it also responded when the effective stimulus was at a parafoveal position and there was an ineffective stimulus for the neuron at the fovea. If there was a response in a parafoveal position, it was possible with the design to show whether the receptive field was symmetric or asymmetric with respect to the fovea, and how large the response was when the effective stimulus was in a parafoveal vs. a foveal position.

The design of the experiments was to use a very wide range of stimuli in the different experiments with different neurons, so that any possible effects of the saliency of any particular stimulus would not bias the results. To this end, we used a set of 141 stimuli during the data collection for these experiments, and any individual stimulus was typically used in only one (63 stimuli) or two (53 stimuli) experiments. Part of the experimental design was also to randomize where the effective and ineffective stimuli were placed on any trial, so that any effects found would be valid independently of any possible interactions between particular pairs of stimuli, but would be valid in the more general case independently for any given position of the effective stimulus of where particular ineffective stimuli were located on a particular trial.

One comparison condition included the same task but with only the effective stimulus in the display on any one trial, at one of the five positions. This condition is termed the 'single' condition. This provided evidence on whether the same neuron had a large receptive field when only one object was present in the scene.

A second comparison condition included the five simultaneously presented stimuli, but now spaced out further from the fovea, to investigate how far from the fovea the effects discovered extend from the fovea. Their centres were removed further from the point of fixation, so that the centre of a parafoveal $7 \times 6^\circ$ object was 16° from the fovea. This condition is termed the 'far' condition. The five simultaneously presented objects could be thought of as forming a

complex scene, but at the same time as allowing the neuronal responses to the defined objects to be clearly measured when they were in different positions with respect to the fovea. For four of the 54 neurons, tests were performed with half-sized objects. In this case, the size of the objects was $4 \times 3^\circ$, and a parafoveal object was located 5° from the fovea in the near condition, and 8° in the far condition. We note here that the results for these four neurons were similar to those obtained with the larger stimuli (with one neuron having asymmetry in the close but not single condition, and one having asymmetry in both the close and the single conditions).

The visual fixation task

Eye position was measured to an accuracy of 0.5° with the scleral search coil technique (Judge *et al.*, 1980), and steady fixation of a position on the monitor screen was ensured by use of a blink version of a visual fixation task in which the fixation spot was blinked off 100 ms before the target stimulus appeared. The eye position was acquired using 1-kHz digitization and storage of new values every 20 ms. Each trial started at -500 ms (with respect to the onset of the test image) with a 500-ms warning tone to allow fixation of the fixation spot, which appeared at the same time. At -100 ms the fixation spot was blinked off so that there was no stimulus on the screen in the 100 ms period immediately preceding the test image (see Fig. 1C). The screen in this period, and at all other times including the inter-stimulus interval, was set at the mean luminance of the test images. At 0 ms, the tone was switched off and the test image was switched on for 1600 ms. At the termination of the test stimulus the fixation spot reappeared, and then after a random interval in the range 150–3350 ms it dimmed, to indicate that licking responses to a tube in front of the mouth would result in the delivery of juice reward. The dimming period was 1000 ms, and after this, the fixation spot was switched off, and reward availability was terminated 500 ms later. The monkey was required to fixate the fixation spot in that if it licked at any time other than when the spot was dimmed, saline instead of fruit juice was delivered from the tube; and in that if the eyes moved by more than a fixed criterion (typically 1° but sometimes 2°) from time 0 until the start of the dimming period, then the trial was aborted. (When a trial aborted, a high-frequency tone sounded for 0.5 s, no reinforcement was available for that trial and the inter-trial interval was lengthened from 8 to 11 s.) It should be noted that this is not a task in which attention either at or away from the test visual stimuli was explicitly rewarded. The reward the monkey received was not contingent on the monkey making any response to the test stimuli. What the fixation task does is allow the visual fixation position of the monkey to be controlled. This then enabled stimuli to be shown in different positions with respect to the fovea. In this way we were able to examine the effects of having more than one visual stimulus present in the visual field at one time.

Recording techniques

The activity of single neurons was recorded with FHC Inc. (Bowdoinham, MA, USA) epoxy-insulated tungsten microelectrodes with impedances of 1–5 M Ω , with the monkey seated in a primate chair using techniques described previously (Rolls *et al.*, 1990; Tovee *et al.*, 1993; Booth & Rolls, 1998). The use of these electrodes allowed good signal-to-noise ratios, typically greater than 3 : 1. The action potentials of single neurons were amplified (Rolls *et al.*, 1979) and acquired with Datawave Enhanced Discovery data acquisition (DataWave, CO,

USA). The isolation of single neurons was assured using Datawave spike cluster cutting off-line.

Procedure

Tracks were made into the IT and the cortex in the superior temporal sulcus (see Fig. 10) and the responses of isolated neurons were measured to a wide variety of small stimuli on the video monitor. These included faces, objects, sine-wave gratings and boundary curvature descriptors (for examples see Fig. 1 and Rolls & Tovee, 1995b). If the neuron responded to some but not other of the stimuli the search continued to find an object to which the neuron had a large response, and four other stimuli to which the neuron had no response, as described above. The experiment was then run with five simultaneously presented objects, until the effective object had been shown in each of the five positions at least eight times. Then the comparison condition with only the effective stimulus being shown in one of the five positions was run, with fixation maintained as always at the centre of the screen. After this, if possible the experiment was repeated in the far condition with the stimuli offset 16° diagonally from the fovea.

Data analysis

The mean response of a neuron when the effective stimulus was in each of the five positions was calculated, and is shown with the standard error of the mean in the figures presented. (Unless specified otherwise, the figures and analyses are based on the responses of the neurons, i.e. the firing rate minus the spontaneous activity.) An ANOVA was performed across the responses of the effective stimulus when in the four parafoveal positions to test whether there was significant asymmetry in the receptive field. The other stimuli, which did not increase the firing rate of the neuron from which recordings were being made, were in random positions on the different trials. The degree of asymmetry of the receptive field for the four parafoveal positions was calculated as (the standard deviation of the firing rates calculated over the four parafoveal positions/the mean firing rate calculated over all five positions). This is termed the asymmetry index. (The form of the index is similar to that for the coefficient of variation, which is defined as the standard deviation/the mean.) The asymmetry index takes the value 0 when there is no difference in firing rate when the effective stimulus is in the different parafoveal positions. (In practice, because there is some random fluctuation in the mean responses, the asymmetry index can become close to zero when the responses at the different parafoveal locations are not significantly different, but is unlikely to reach zero.) When calculating this index, the spontaneous firing rate was subtracted so that the neuronal response was measured, as elsewhere in this paper. (If the subtraction resulted in a negative value, then all the firing rates were corrected by setting the lowest value to zero.) Trials on which fixation was not maintained were excluded from the analysis. The firing rate was measured in a 500 ms period starting 100 ms after the visual array appeared, as this was a typical response latency for anterior IT neurons described here and reflects retinal delays as well as the cortical processing in each of the stages from the primary visual cortex to the inferior temporal cortex, which is in the order of 15 ms per stage (Panzeri *et al.*, 2001; Rolls & Deco, 2002).

Calibration trials for the scleral eye position recording system were performed. To obtain calibration data, the output of the eye position monitoring system was measured while the monkey performed a visual fixation task with fixation points in a five-position array (upper left, upper right, screen centre, lower left, and lower right).

Recording sites

X-radiographs 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 & Rolls, 1991). Reconstructions of the tracks were made from serial 50 μm histological sections using the positions of the microlesions and the reference electrodes, and from the corresponding X-ray coordinates of all recorded cells, to determine the locations of all the cells.

Results

Recordings were made from well-isolated IT neurons in two macaques. In the course of these recordings, it was possible to find a reasonable number of neurons that responded well and with selectivity to some of the small test images of objects, faces, etc., that were available, with the proportions of different types of responsive cell similar to those we have reported previously (Baylis *et al.*, 1987). To perform the experiments described here, it was necessary to find visually responsive neurons in the temporal cortical visual areas, and to select a set of five stimuli, one of which increased the firing rate of the neuron being recorded, and the others of which had only small effects on the firing rate, i.e. where no significant effect was found in the peristimulus time histogram. The total number of neurons recorded in the tracks was 187, and of these 145 gave an indication of visual responsiveness. The experiments described here were completed on 54 neurons. These 54 neurons had significant differences in their firing rates between the five positions (five parafoveal and one foveal) in any of the three conditions tested, namely five stimuli simultaneously present in the close (10°) configuration, five stimuli simultaneously present in the far (16°) configuration, and the effective stimulus only being presented in one of the five positions (the single condition). This set could thus include neurons with symmetric receptive fields, but where the firing rate when the effective stimulus was in the four parafoveal positions was smaller than when the effective stimulus was at the fovea. We note that for the 54 neurons studied here, there was no selection criterion based on whether the responses in the parafoveal locations were symmetric vs. asymmetric. Fifteen neurons additional to the 54 had uniform responses in that there was no difference in the firing rate that depended in the close condition on where the effective stimulus was shown (i.e. in which of the five positions the effective stimulus was located). Thus, 15 of the neurons examined with five stimuli present had symmetric and large receptive fields, in that the responses were not different when the effective stimulus was in any of the parafoveal and foveal positions. Four additional neurons could respond if one visual stimulus was on the screen, but had no response if that stimulus was presented simultaneously with four other stimuli in the close condition. The remaining neurons of the 145 could not be held sufficiently long for the experiments to be completed.

First, the results show that most (51) of the 54 neurons had a significant response to the effective stimulus when it was presented at the fovea simultaneously with four other stimuli in the parafovea. Indeed, with the effective stimulus at the fovea and four stimuli in parafoveal locations, the mean (\pm SD) response was 10.8 ± 8.5 spikes/s, compared with 13.4 ± 8.3 spikes/s when only the effective stimulus was present and at the fovea. (These response rates are for the neurons tested under both conditions, and are not significantly different, paired $t = 0.114$.) (The average spontaneous firing rate of the neurons was 13.9 spikes/s, and thus the mean firing rate across the

whole population of neurons to the effective stimulus when only the effective stimulus was present was 27.3 spikes/s.) Thus, the receptive fields of the majority of the 54 neurons included the fovea even when five stimuli were simultaneously present in the receptive field.

Second, when multiple objects are present in a scene, 87% of the 54 IT neurons responded to their effective stimulus at least partly when it was centred 10° from the fovea. Moreover, eight IT neurons had a significantly better response to an object when it was located in one or more of the parafoveal positions than when it was at the fovea ($P < 0.05$ in a *post-hoc* least-significant difference test following an ANOVA on the firing rates in the foveal and four parafoveal positions). In the remainder of the paper, we focus on how these 54 neurons responded when stimuli were present in the parafoveal positions, in order to examine how IT neurons provide evidence about stimuli that are not at the fovea, especially when multiple objects are present simultaneously in a scene.

The responses of one neuron when tested with five stimuli present in the close condition (with the parafoveal stimuli 10° from the centre) are shown in Fig. 1A, and the responses of the same neuron when only the effective stimulus was presented are shown in Fig. 1B. Figure 1C shows an eye position recording to show that fixation at the centre of the screen was maintained throughout the trial. First, it is of interest to note that even in the crowded scene with five simultaneously presented stimuli (Fig. 1A), the neuron did respond to the effective stimulus not only when it was at the fovea, but also when it was in at least some of the parafoveal locations. The implication is that some IT neurons can respond to stimuli which are not at the fovea. To quantify this, we calculated a mean parafoveal index as the mean response to the effective stimulus in the four parafoveal positions with five stimuli present divided by the response when the effective stimulus was at the fovea with five stimuli present. (The response is the firing rate minus the spontaneous rate.) For this neuron, the mean parafoveal index was 0.51. We also calculated a maximum parafoveal index as the response to the effective stimulus in its most effective parafoveal position with five stimuli present divided by the response when the effective stimulus was at the fovea with five stimuli present. For this neuron, the maximum parafoveal index was 1.16. Second, Fig. 1A shows that the neuron had statistically significant asymmetry in its parafoveal responses when five stimuli were present, as shown by an ANOVA performed over the four parafoveal positions of the effective stimulus for the neuron ($F_{3,33} = 9.7$, $P < 0.001$). When only one stimulus was present, there was no significant asymmetry ($F_{3,37} = 1.2$, ns) (data shown in Fig. 1B). An asymmetry index was calculated as the standard deviation of the neuronal responses calculated over the four parafoveal positions divided by the mean neuronal responses calculated over all five positions. The asymmetry index for this neuron when tested across the four parafoveal positions as described with five stimuli simultaneously present was 0.77 (data shown in Fig. 1A), and when tested with only one stimulus present was 0.17 (data shown in Fig. 1B).

The responses of two more neurons to illustrate the types of asymmetry found when multiple objects are presented simultaneously in the scene are shown in Fig. 2. The responses of the neuron shown in Fig. 2B occurred mainly when the effective stimulus was at the fovea, and much less for any parafoveal position. The mean parafoveal index was 0.18, and the maximum parafoveal index was 0.29 for this neuron, with five stimuli present. [There was a small difference in the firing rate for the different parafoveal positions of the effective stimulus, which was just significant ($F_{3,84} = 3.15$, $P < 0.03$), but the four peripheral responses were very close to the spontaneous firing rate.] The asymmetry index for the neuron in Fig. 2B when tested with five stimuli simultaneously present was 0.40.

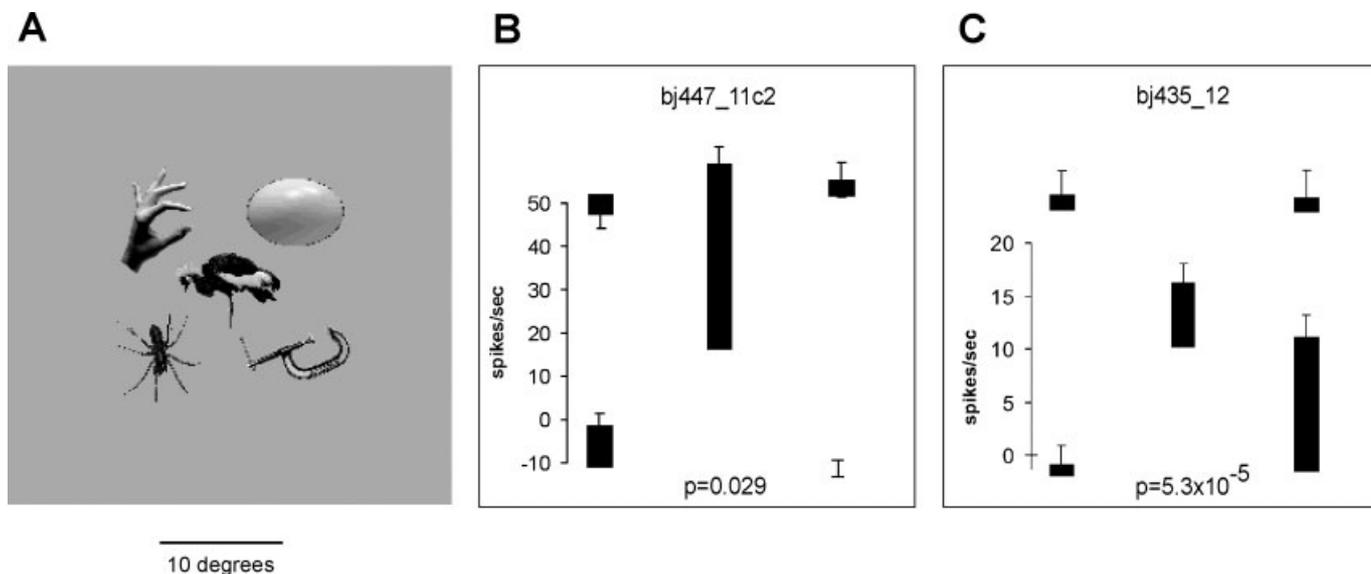


FIG. 2. The responses of two more neurons (B and C) to illustrate the types of asymmetry found when multiple objects are presented simultaneously in the scene (see text). Conventions as in Fig. 1. The responses of the neurons are shown, i.e. the spontaneous firing rate was subtracted.

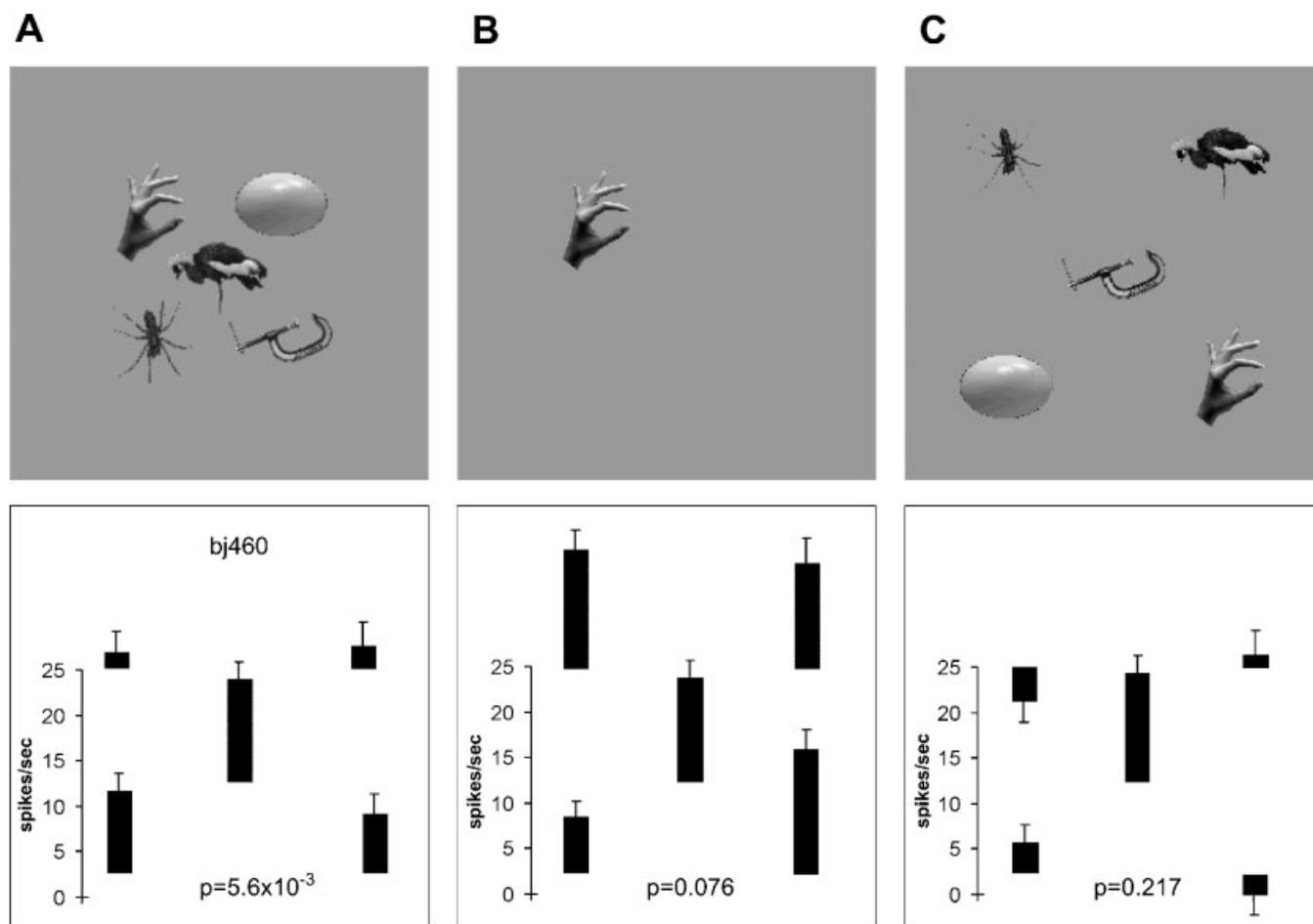


FIG. 3. The responses of a neuron to an effective stimulus tested under all three conditions (A) in the close condition (five stimuli presented simultaneously at the fovea and the four positions 10° from the fovea) (B) in the single condition (the effective stimulus presented alone in the fovea or one of the four parafoveal positions 10° from the fovea), and (C) in the far condition (five stimuli presented simultaneously at the fovea and the four positions 16° from the fovea). Conventions as in Fig. 1.

The responses of the neuron shown in Fig. 2C occurred best when the effective stimulus was to the lower right of the fovea, with little response in any other parafoveal position. There was some response when the effective stimulus was at the fovea. [We note that the large responses of the neuron when the effective stimulus was shown in the lower right occurred as an average result across a very large number of combinations of the other four (ineffective) stimuli in the other four positions, so that the asymmetry shown is what would be expected to occur with a random set of objects in random positions with respect to the fovea. The asymmetry described in this paper is thus a generic type of asymmetry which holds on average when an effective stimulus is presented in a given position independently of the particular positions of the ineffective stimuli for a neuron.] The mean parafoveal index was 0.65, and the maximum parafoveal index was 2.05 for this neuron, with five stimuli present. The neuron had statistically significant asymmetry in its parafoveal responses when five stimuli were present, as shown by an ANOVA performed over the four parafoveal positions of the effective stimulus for the neuron ($F_{3,75} = 8.7$, $P < 0.0002$). The asymmetry index for the neuron in Fig. 2C when tested with five stimuli simultaneously present was 1.29. Although most IT neurons had their best response to an effective stimulus when it was at the fovea, eight of the 54 IT neurons recorded in the close condition had a significantly better response to an effective stimulus when it was located in one or more of the parafoveal positions.

The responses of a neuron tested with the stimuli 16° from the fovea (the far condition) as well as in the other conditions are shown in

Fig. 3. Figure 3A shows that with five stimuli presented simultaneously and the standard 10° separation between the centres of the stimuli, the neuron had asymmetric responses to the different parafoveal locations of the effective stimulus ($F_{3,77} = 4.55$, $P < 0.006$), with some responses when the effective stimulus was below the fovea (especially below left). The asymmetry index was 0.54. The mean parafoveal index with all five stimuli present was 0.64, and the maximum parafoveal index was 0.79. Figure 3B shows the responses to the effective stimulus alone for each position. There was much more translation invariance ($F_{3,36} = 2.51$, ns), and the asymmetry index was lower, 0.37. (As noted in the Methods where the asymmetry index is defined, its value is expected to be low, but not to reach 0 because of the random fluctuation in the mean of the responses at any one location, even when there is no significant asymmetry as shown by the ANOVA.) The mean parafoveal index with one stimulus present was 0.98, and the maximum parafoveal index was 1.25. Figure 3C shows that with five stimuli presented simultaneously and 16° separation between the centres of the stimuli, the neuron had only minor responses to the effective stimuli in the parafoveal position. The mean parafoveal index with all five stimuli present in the far (16°) condition was 0.23, and the maximum parafoveal index was 0.29. There were no asymmetric responses to the different parafoveal locations of the effective stimulus ($F_{3,69} = 1.5$, ns).

For the whole populations of neurons tested in each condition, the mean parafoveal index of each neuron is shown in Fig. 4A. This shows that most neurons had a mean parafoveal index close to 1 when tested with one stimulus present (median 0.91, mean 0.88). Thus, on

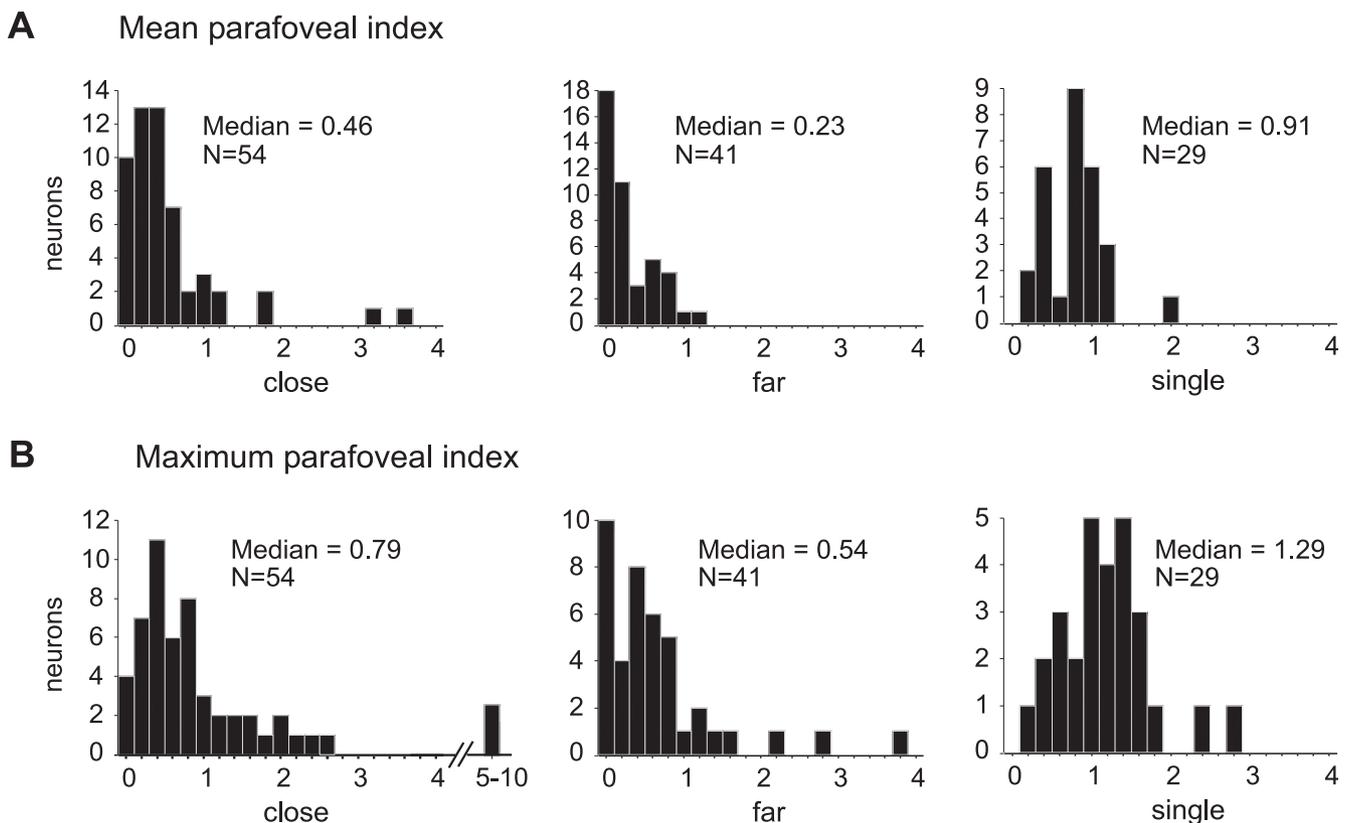


FIG. 4. The mean parafoveal index and the maximum parafoveal index, for all the neurons tested in each of the three conditions, close (five simultaneously presented stimuli centred 10° apart), far (five simultaneously presented stimuli centred 16° apart) and single (one stimulus presented at a time in the same five positions centred 10° apart). The mean parafoveal index of each neuron was the mean response with the effective stimulus in the four parafoveal positions divided by the response with the effective stimulus at the fovea. The maximum parafoveal index of each neuron was the maximum firing rate value in any parafoveal position divided by the firing rate with the effective stimulus at the fovea.

average, with one stimulus present, most neurons responded as well to a parafoveal stimulus as to a foveal stimulus. Figure 4A also shows that when tested with five stimuli simultaneously present in the close (10°) configuration, the mean parafoveal index for most cells was less than 1 (median 0.46, mean 0.64), indicating that in this crowded scene condition, most of the neurons had on average a smaller response in the parafoveal positions with five stimuli present. (The difference between the single and close conditions was significant at $P < 0.0002$, Mann–Whitney test.) The mean parafoveal index was even lower when the simultaneously present parafoveal stimuli were in the far (16°) parafoveal position (median 0.23, mean 0.36). (The difference between the single and far conditions was significant at $P = 10^{-7}$, Mann–Whitney test. The difference between the close and the far conditions was significant at $P = 0.01$, Mann–Whitney test.)

The corresponding maximum parafoveal indices are shown in Fig. 4B. This shows that in the close (10°) condition, typically presentation of the effective stimulus at the most effective parafoveal location produces a response that for most neurons is quite similar to the response when the effective stimulus is at the fovea in the five-stimulus condition (median 0.79, mean 1.31). (Statistically, the maximum parafoveal index in the close condition is significantly larger than the mean parafoveal index in the close condition, Wilcoxon t , $P < 10^{-9}$.) This is consistent with the concept that while a neuron can respond to the effective stimulus in at least one parafoveal position, on average the response in the parafoveal locations is reduced in the close condition (as shown in Fig. 4A). This thus reflects asymmetry in the receptive fields when five stimuli are simultaneously present in the close condition. (While any variability in the neuronal responses will result in the maximum parafoveal index being greater than the mean parafoveal index, the interpretation of particular values for both is evident from Figs 1–3 and the values for the indices given in the text.) The combination of the facts that in the close condition the mean parafoveal index was much less than 1 (median value 0.46, Fig. 4A) whereas in the single condition it was close to 1 (median value 0.91, Fig. 4A), and that in the close condition the neurons could still fire fast in some locations as shown by the high values of the maximum parafoveal index (median value 0.79, Fig. 4B), is an indication of considerable asymmetry in the close condition. In the far condition, even the maximum parafoveal indices are reduced below 1 for most of the neurons, indicating that most of the neurons do not have large responses to the stimuli in any parafoveal position when the parafoveal stimuli are in the far configuration (median 0.54, mean 0.75). (Statistically, the maximum parafoveal index is smaller in the far than the single condition, Mann–Whitney $P = 1.5 \times 10^{-5}$; and smaller in the far than the close condition, Mann–Whitney $P = 0.036$.) With just one stimulus present, the maximum parafoveal index had a median value of 1.33, and a mean of 1.30, indicating that the neurons responded with one stimulus present just as well (and by chance even a little more given that the highest of four observations was taken than to the effective stimulus at the fovea).

Further evidence on the types of asymmetry shown in this population of neurons under the close and single testing conditions is shown by the examples in Fig. 5 of five neurons. The neurons show considerable asymmetry in the close condition (as shown by the

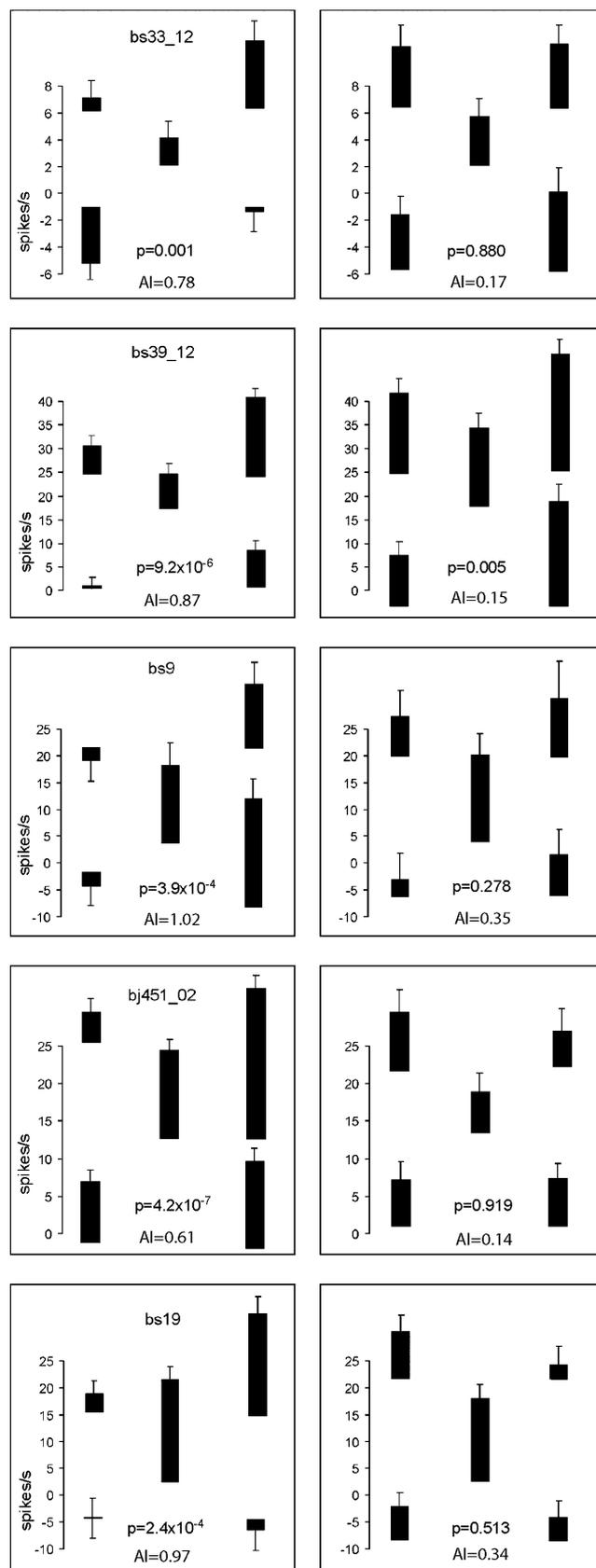
Close (10°)Single (10°)

FIG. 5. The responses (mean \pm SEM) of five neurons when tested in the simultaneous close (10°) and single conditions. The responses are shown when the effective stimulus for the neuron was in the relevant position. The experiment number is shown on the top, and the asymmetry index (AI) is also shown (see text). The results of the ANOVAs calculated across the four parafoveal conditions are also indicated by the P -values.

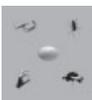
	Total number of neurons tested	Number of neurons with symmetric receptive fields	Number of neurons with asymmetric receptive fields
	54	19	35 (65%)
	29	23	6 (21%)
	41	30	11 (27%)

FIG. 6. Summary of the numbers of neurons with asymmetry evident in the receptive fields under different testing conditions. The measure of asymmetry was a significant difference in the responses between the four different parafoveal positions in which the effective stimulus was shown. Top row: the close condition (five simultaneously presented stimuli centred 10° apart). Middle row: the single condition (the effective stimulus only was presented in each of the five positions which were 10° apart). Bottom row: the far condition (five simultaneously presented stimuli centred 16° apart).

ANOVA calculated over the four parafoveal positions), and this is not just due to any possible contralateral biasing of the receptive fields, because most of these neurons show differences in their responses between the upper and lower parafoveal positions (as shown by *post-hoc* tests to the ANOVAs). In the single condition, most of the neurons have responses in all the parafoveal positions, and the responses are relatively similar for the different parafoveal positions, as shown by the ANOVAs calculated over the four parafoveal positions.

To compare the asymmetries in the different testing conditions (close, single, far), we first performed ANOVAs on the firing rates in the four parafoveal positions for all neurons tested, to show whether there were significant differences in the firing rates in the parafoveal positions. The results are shown in Fig. 6. There are relatively many neurons with asymmetric receptive fields in the close compared with the one-stimulus condition, and this was confirmed by a chi-squared test ($\chi^2 = 13.0$, d.f. = 1, $P < 0.001$). There are also relatively many neurons with asymmetric receptive fields in the close compared with the far condition, and this was confirmed by a chi-squared test ($\chi^2 = 12.0$, d.f. = 1, $P < 0.001$). We checked that these results still held when ANOVAs were performed just on the subset of cells tested in both the close and the single conditions with the same numbers of trials in each condition, and found that 9/29 tested had significant asymmetries in the close condition, and two of the same 29 tested in the single condition had significant asymmetries ($P < 0.01$) as shown by the ANOVAs. [A McNemar test for the significance of changes (Siegal, 1956), which explicitly tests whether different numbers of the same neurons changed from asymmetric fields in the close to symmetric in the single condition, showed $\chi^2 = 4.5$, d.f. = 1, $P < 0.05$, indicating that the same cells were more likely to have asymmetric fields in the close than the single condition. This McNemar test is a control for the situation that some of the neurons, though not all, included in the chi-squared tests above were tested in both conditions.] The proportion of neurons with asymmetric receptive fields in the far condition was 27%.

We note that the values of the mean parafoveal indices shown in Fig. 4 are reduced to approximately half in the close condition with simultaneously presented stimuli, providing an estimate of the

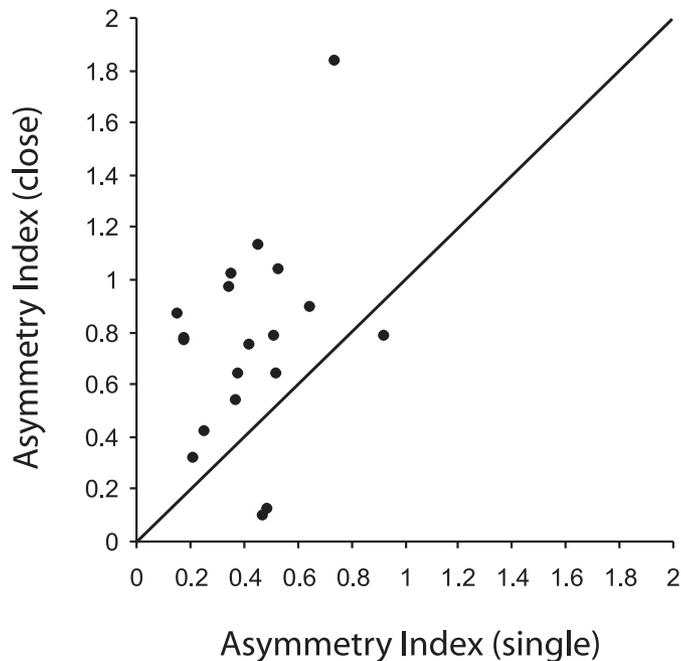


FIG. 7. The asymmetry indices for the neurons when the five simultaneously presented stimuli were in the close condition (with the parafoveal stimuli 10° from the fovea), and when one stimulus was present (single). The asymmetry index is the standard deviation of the firing rates across the four parafoveal positions divided by the mean firing rate calculated over all five positions. Each point is for one neuron. The line at 45° shows where the points would lie if their asymmetry index was the same in the close and single conditions.

receptive field size with five simultaneously presented stimuli of 20° in diameter. This compares with estimates of approximately 20° in diameter when two stimuli are shown in a complex natural scene (Rolls *et al.*, 2003). The present results are thus in agreement with the earlier results on receptive field size, in that having multiple objects in a scene (as in this paper), or two objects in a complex natural scene (Rolls *et al.*, 2003), thus does reduce receptive field size to approximately 20°, compared with the 70° when a single stimulus is shown in a blank screen (Rolls *et al.*, 2003).

The asymmetry indices for the population of neurons when the five stimuli were separated by 10° are shown in Fig. 7 (close condition) for each neuron. (All neurons tested in both the close and the single conditions with a significant difference in the firing rates across the four parafoveal positions are included in this figure.) The median asymmetry index was 0.78 (mean 0.76). For comparison, the asymmetry indices for the same neurons when only one stimulus was presented are also shown in Fig. 7 (single). The median asymmetry index was 0.42 (mean 0.42). (As noted in the Methods where the asymmetry index is defined, its value is expected to be low, but not to reach 0 because of the random fluctuation in the mean of the responses at any one location, even when there is no significant asymmetry as shown by the ANOVA.) A Wilcoxon *t*-test for the 19 neurons tested in both conditions in which there was statistically significant asymmetry in either the close or the single condition showed $t = 24$, $n = 19$, $P = 0.004$. Thus, these neurons had more asymmetry by this measure in their parafoveal responses when multiple objects were present in the receptive fields, and less asymmetry (i.e. more translation invariance) when only one stimulus was presented on a blank screen.

Different neurons had their asymmetry in different positions with respect to the fovea. This is shown in Fig. 8, which shows the mean

Average firing rate for each parafoveal position
1= top ipsilateral, 4 = bottom contralateral

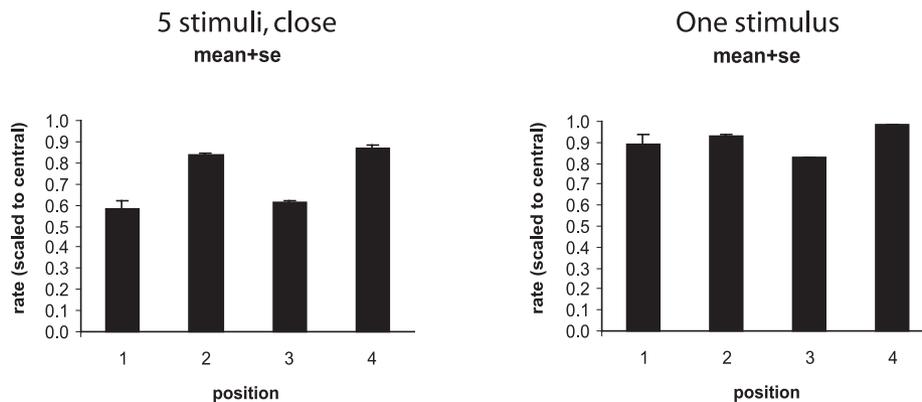


FIG. 8. The mean firing rate across neurons with the effective stimulus at each parafoveal position expressed as a fraction of the firing rate when the effective stimulus was at the fovea (\pm SEM). Position 1 is upper ipsilateral, position 2 is upper contralateral, position 3 is lower ipsilateral, and position 4 is lower contralateral. Left: the close (10°) condition. Right: with a single stimulus presented. The firing rate for each neuron was scaled to the firing rate at the fovea before the mean value shown was calculated.

firing rate across neurons for each parafoveal position expressed as a fraction of the firing rate when the effective stimulus was at the fovea. It is evident that all parafoveal positions are represented among the population of neurons, although there is some contralateral bias (position 2 is upper contralateral, 4 is lower contralateral), but a similar average response across the neuronal population for the upper vs. the lower parafoveal region of the visual field. Furthermore, the asymmetries found did not reflect just a contralateral bias, in that 25 of the 35 neurons with asymmetric receptive fields had significant differences in the firing rates when the effective stimuli were in an upper vs. a lower position on the same side ($P < 0.05$ in a fully corrected *post-hoc* test to the ANOVA). It is shown (left) that when all five stimuli are present, the firing rates are on average lower in the parafoveal positions than when only one stimulus is present (right), and that with only one stimulus present, the firing is very similar for the different parafoveal positions.

It was of interest that the latencies and time courses of the neuronal responses were similar in the single and multiple stimulus (close) conditions. An example is shown in Fig. 9 of the time course of the responses of a neuron in the single and close conditions. For all 19 neurons shown in Fig. 7 the latencies of any responses that were present were similar for the single and close conditions. Furthermore, in 17/19 cases the latencies of the responses were similar independently of where the effective stimulus was in the close or single condition.

Recording tracks were made over an extensive portion of the IT, 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. Illustrations of the recording sites of the cells reported here are provided in Fig. 10. As can be seen, the cells are distributed in a region that extends from the gyrus lateral to the middle temporal sulcus, to the lower bank of the superior temporal sulcus, and the area of cortex investigated is indicated by the shaded bounding box on the upper right coronal section.

Discussion

The results show that most (51/54) of the neurons had significant responses to the effective stimulus when it was presented at the fovea

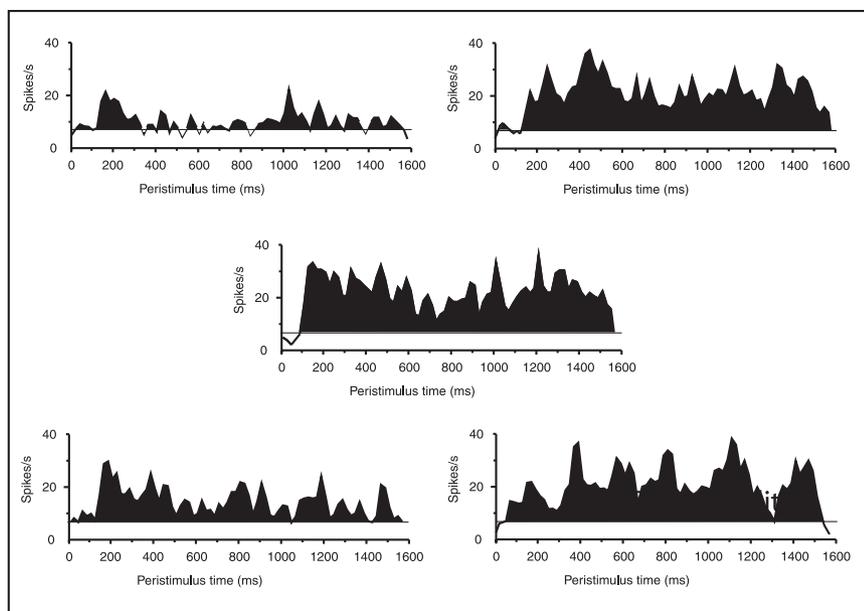
simultaneously with four other stimuli in the parafovea, and that the firing rates of the neurons were not significantly different when the effective stimulus was at the fovea, and there were four stimuli in the parafovea or no stimuli in the parafovea. Thus, the receptive fields of most of the 54 neurons included the fovea even when five stimuli were simultaneously present in the receptive field.

When multiple objects are present in a scene, 87% of IT neurons responded to their effective stimulus at least partly when it was centred 10° from the fovea. This is reflected in the parafoveal indices shown in Fig. 4. Thus, the output of the inferior temporal visual cortex is about more than what is at the fovea, and many IT neurons can respond to the effective stimulus when it is in at least some parafoveal positions.

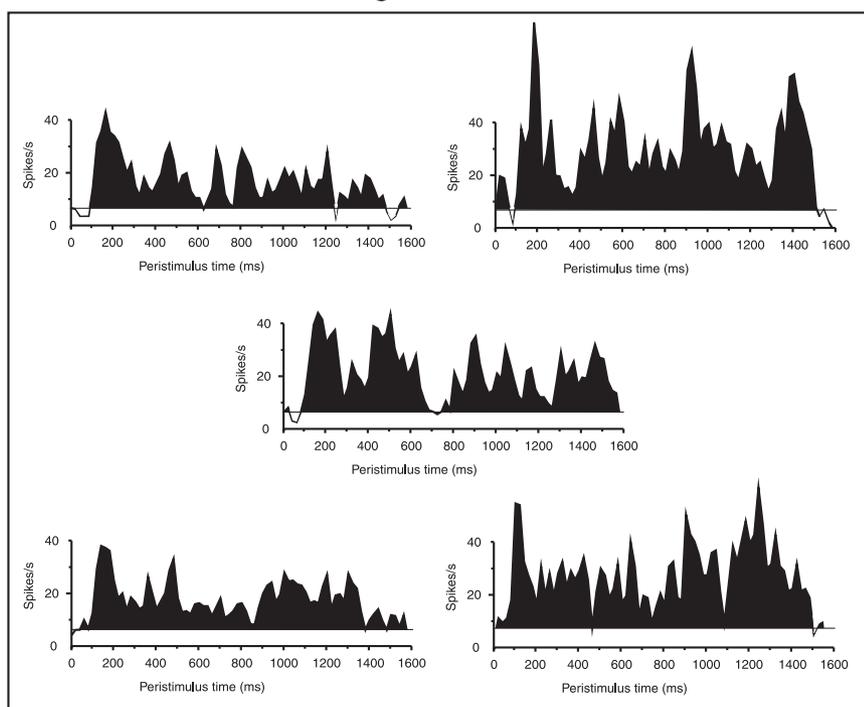
When multiple objects are present in a scene, 65% of IT neurons responded asymmetrically with respect to the parafoveal position of the effective stimulus, as shown by significant differences in the firing rates when the effective stimulus was at different parafoveal positions (Fig. 6). Furthermore, Fig. 4 shows that in the close condition (left) the maximum parafoveal index (median 0.79) is higher than the mean parafoveal index (median 0.46). Additionally, the mean asymmetry index in the close condition had a median value of 0.78 (Fig. 7).

The same neurons have large, more symmetric receptive fields when only one object is present in the scene. This is shown by the fact that with just one stimulus present, only 21% of the neurons had significantly different firing rates in the different parafoveal positions (Fig. 6). In addition, the asymmetry indices had significantly lower values, indicating lower asymmetry in the single than in the close condition, as shown in Fig. 7, and the mean parafoveal index values were higher (Fig. 4, single). The implication of the comparison between the greater asymmetry in the close simultaneous (10°) than single conditions is that interactions between stimuli within the receptive field increase the asymmetry that is found. Thus, in natural scene viewing conditions, when many objects are likely to be present, the asymmetry is likely to be significant, and the neurons will convey evidence not only about what is at the fovea, but also about what is present in the periphery, and where it is present. Nevertheless, the asymmetry found when multiple objects are presented simultaneously may be developed from some asymmetry that is present when only one stimulus is being presented. This is illustrated in Figs 5 and 7, and quantified by the fact that even in the single stimulus condition, the

A. Close condition



B. Single condition



C. Ineffective stimulus, single condition

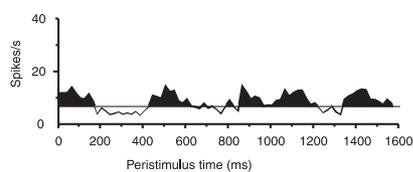


FIG. 9. Peristimulus time histograms to show the time course of the responses of a neuron in the (A) close, 10° , and (B) single conditions. Separate histograms are shown for presentation of the effective stimulus at the fovea and at each of the four parafoveal positions. The stimuli were presented at time 0. (C) The lack of response to an ineffective stimulus averaged across all five positions. The histograms are shaded in the part higher than the spontaneous firing rate of the neuron which was 7 spikes/s.

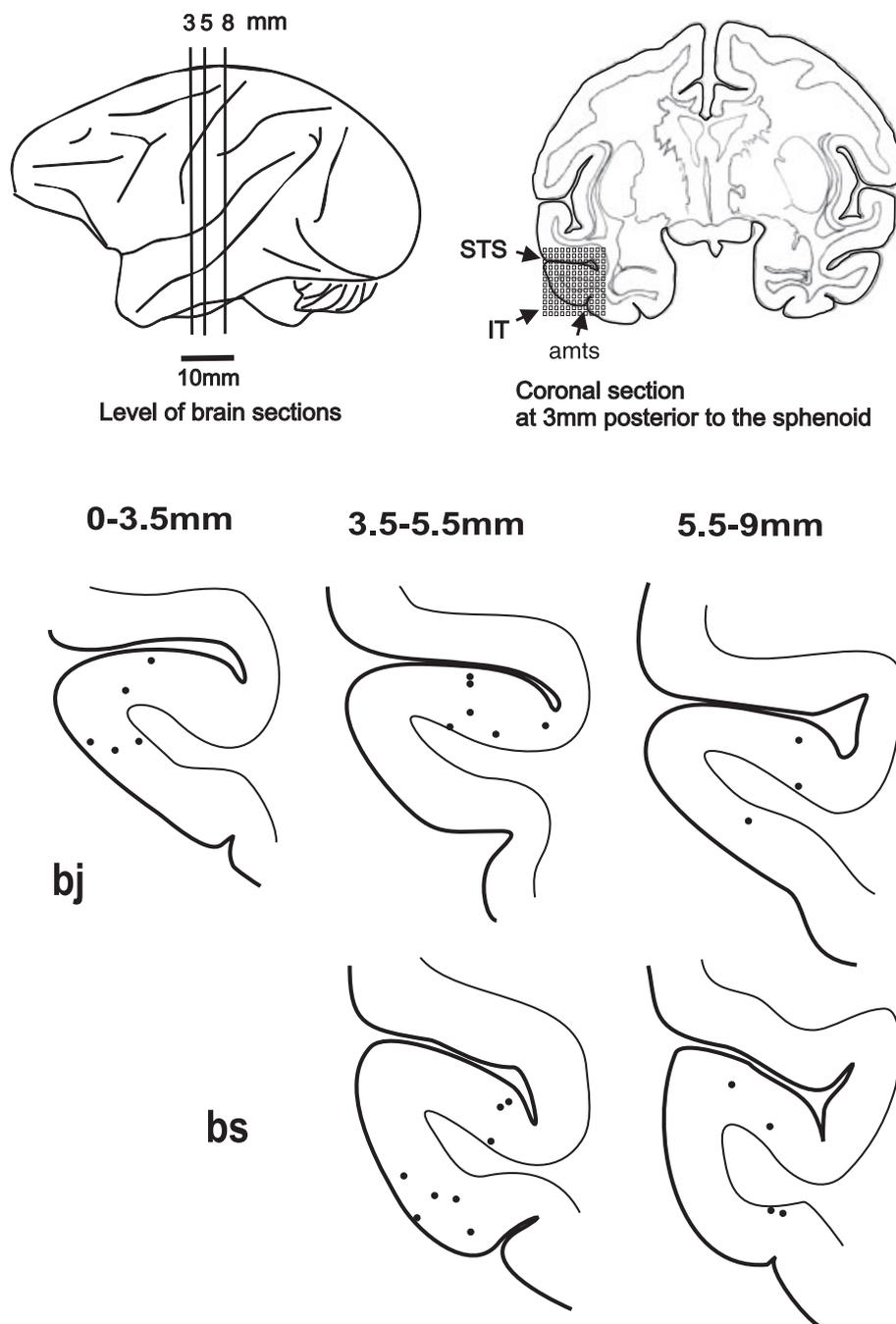


FIG. 10. Reconstructed histological coronal sections showing (filled circles) the sites at which the neurons analysed in this paper were recorded in macaques bj and bs. Numbers above the sections indicate the distance (in mm) posterior to the sphenoid bone reference point (which is at approximately the anterior–posterior level of the anterior commissure), and these positions are further illustrated in the upper left of the figure in lateral view. A full coronal section is illustrated at the top right of the figure, 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).

mean value of the asymmetry index was 0.42. (When the ANOVA showed no significant asymmetry, values of the asymmetry index were typically in the range 0.1–0.35 due to the statistical variations in the firing rate, as explained in the Methods.) Further evidence that the receptive fields of IT neurons can be small and not centred at the fovea when just one stimulus is presented has been found with small (0.6°) stimuli in macaques trained to discriminate between stimuli with distractors present (DiCarlo & Maunsell, 2003). Irregularities in the receptive fields of IT neurons have also been found when one stimulus

is presented by Op De Beeck & Vogels (2000). We emphasize that here we investigated responsiveness and asymmetry well away from the fovea, with stimuli located 10° into the parafovea, in order to understand whole scene perception.

The mechanism by which the asymmetry is produced in the multiple simultaneous stimulus condition is probably related in part to feedback inhibition, which can tend to weight what is at the fovea due to the greater magnification factor at the fovea (Rolls & Deco, 2002; Trappenberg *et al.*, 2002; Deco & Rolls, 2004). If this competition by

shunting inhibition is superimposed on some inherent asymmetry in the receptive fields of IT neurons, which is clearly present as shown by the finding that even with one stimulus present 21% of IT neurons responded significantly differently to stimuli in different positions (Fig. 6) (see also Op De Beeck & Vogels, 2000; DiCarlo & Maunsell, 2003), then this could produce the greater asymmetry evident when five stimuli are located at or close to the fovea, as studied here, or in natural scenes.

When multiple objects are present in a scene, these neurons are less likely to respond to their effective stimulus when it is 16° from the fovea: the median value of the mean parafoveal index was 0.23 (Fig. 4). Furthermore, the mean response of the neurons to the effective stimulus in its parafoveal positions in the far condition was 6.0 spikes/s (± 1.2 , SEM), compared with a mean when at the fovea of 14.7 spikes/s (± 2.6 , SEM). Nevertheless, there was sometimes some response in the 16° (far) parafoveal position (see Fig. 3C for example), and if there was a response, there was sometimes asymmetry, as shown by the finding that the median value of the maximum parafoveal index was 0.54 (Fig. 4).

The type of encoding uncovered combines considerable stimulus specificity, with some information about the location of the stimulus. The encoded information is that the effective stimulus for that neuron is present, and is in general either at the fovea or is in one or two of a few parafoveal positions. Across an ensemble of such neurons, information that a given stimulus is present is provided, and because each neuron responding to an effective stimulus may have an asymmetric parafoveal response, the effective stimulus is exactly at the fovea, or is displaced a few degrees in a direction encoded by the particular subset of asymmetric neurons tuned to that stimulus that is active. The result is that in a single glance, of say 100 ms, insufficient time for an eye movement, the IT provides information in a cluttered scene with multiple objects about which object is at the fovea, and about which objects are in parafoveal locations (centred 10° away from the fovea). It is notable that the encoding is in labelled line form, with no need for syntactic binding (Singer, 1999; Rolls & Deco, 2002), in that different IT neurons are active for the effective stimulus when it is at different parafoveal positions. The type of encoding revealed here thus shows a solution that the visual system provides to the representation of multiple objects in a scene, which is an important issue in hierarchically convergent object recognition systems with distributed representations (Mozer, 1991). Indeed, the type of encoding described in this paper could account for how we are able to see two versions of the same object at different positions in a scene.

How then should we understand the encoding by IT in complex natural scenes? One key conclusion is that whereas the receptive fields of many IT neurons are very large, typically 70° in diameter, with blank backgrounds, the receptive fields shrink in complex natural scenes down to be close to the size of objects, and include or are close to the fovea (Rolls *et al.*, 2003). This helps to minimize the binding problem in complex scenes, in that just a set of features relatively close to the fovea need to be bound together into an object. It also helps succeeding areas to read the information for memory or action in a simple way, as the information is primarily about the object at the fovea (Rolls & Deco, 2002; Rolls *et al.*, 2003). The extension to these concepts provided here found as a result of presenting five stimuli in defined positions with respect to the fovea is that the output of many IT neurons provides in complex scenes information not only about the object at the fovea, but also some information about other objects close to the fovea, located 10° from the fovea. Given that the total firing for the object at the fovea will dominate the IT output (given the typically smaller firing in parafoveal locations measured by the mean

parafoveal index of 0.46 in this study), the parafoveal activity will not dominate the use of information about which object is at the fovea in succeeding areas. However, as some information about objects at parafoveal locations is provided when a set of objects is simultaneously presented, this could be useful in whole scene perception, as there is some evidence even from a single glance about where other objects are located with respect to the fovea.

These results provide evidence that the IT output may be appropriate for the selection of the reinforcing properties of what is at the fovea (Rolls, 2005), as well as selection of the target for action, by virtue of the fact that its receptive fields typically become smaller in complex scenes, but still emphasize what is at the fovea. At the same time, the receptive fields by including some information about where particular objects are in the parafovea may help whole scene processing.

Acknowledgements

This research was supported by the BBSRC, The Wellcome Trust and the Medical Research Council.

Abbreviation

IT, inferior temporal cortex.

References

- Adler, F.H. (1959) *Physiology of the Eye*. The C. V. Mosby Co., St Louis.
- Baddeley, R.J., Abbott, L.F., Booth, M.J.A., Sengpiel, F., Freeman, T., Wakeman, E.A. & Rolls, E.T. (1997) Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. Lond. B*, **264**, 1775–1783.
- Baylis, G.C., Rolls, E.T. & Leonard, C.M. (1987) Functional subdivisions of the temporal lobe neocortex. *J. Neurosci.*, **7**, 330–342.
- Biederman, I. (1972) Perceiving real-world scenes. *Science*, **177**, 77–80.
- Biederman, I., Rabinowitz, J., Glass, A.L. & Stacy, E.W. Jr (1974) On the information extracted from a glance at a scene. *J. Exp. Psych.*, **103**, 597–600.
- Booth, M.C.A. & Rolls, E.T. (1998) View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb. Cortex*, **8**, 510–523.
- Chelazzi, L. (1998) Serial attention mechanisms in visual search: a critical look at the evidence. *Psychol. Res.*, **62**, 195–219.
- Chelazzi, L. & Corbetta, M. (2000) Cortical mechanisms of visuospatial attention in the primate brain. In Gazzaniga, M.S. (Ed.), *The New Cognitive Neurosciences*. MIT Press, Cambridge, MA, pp. 667–686.
- Chelazzi, L., Miller, E.K., Duncan, J. & Desimone, R.E. (1993) A neural basis for visual search in inferior temporal cortex. *Nature*, **363**, 345–347.
- Deco, G. & Rolls, E.T. (2004) A neurodynamical cortical model of visual attention and invariant object recognition. *Vis. Res.*, **44**, 621–644.
- Desimone, R. (1991) Face-selective cells in the temporal cortex monkeys. *J. Cogn. Neurosci.*, **3**, 1–8.
- Desimone, R. & Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.*, **18**, 193–222.
- DiCarlo, J.J. & Maunsell, J.H. (2003) Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *J. Neurophysiol.*, **89**, 3264–3278.
- Feigenbaum, J.D. & Rolls, E.T. (1991) Allocentric and egocentric spatial information processing in the hippocampal formation of the behaving primate. *Psychobiology*, **19**, 21–40.
- Gross, C.G., Desimone, R., Albright, T.D. & Schwartz, E.L. (1985) Inferior temporal cortex and pattern recognition. *Exp. Brain Res.*, **11**, 179–201.
- Gross, C.G., Rocha Miranda, C.E. & Bender, D.B. (1972) Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol.*, **35**, 96–111.
- Hubel, D.H. & Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in cat's visual cortex. *J. Physiol.*, **160**, 106–154.
- Judge, S.J., Richmond, B.J. & Chu, F.C. (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. *Vis. Res.*, **20**, 535–538.

- Miller, E.K., Gochin, P.M. & Gross, C.G. (1993) Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Res.*, **616**, 25–29.
- Missall, M., Vogels, R., Chao-Yi, L. & Orban, G.A. (1999) Shape interactions in inferior temporal neurons. *J. Neurophysiol.*, **82**, 131–142.
- Mozier, M. (1991) *The Perception of Multiple Objects: a Connectionist Approach*. MIT Press, Cambridge, MA.
- Op De Beeck, H. & Vogels, R. (2000) Spatial sensitivity of macaque inferior temporal neurons. *J. Comp. Neurol.*, **426**, 505–518.
- Palmer, S.E. (1999) *Vision Science: Photons to Phenomenology*. MIT Press, Cambridge, MA.
- Panzeri, S., Rolls, E.T., Battaglia, F. & Lavis, R. (2001) Speed of feedforward and recurrent processing in multilayer networks of integrate-and-fire neurons. *Network: Computation Neural Systems*, **12**, 423–440.
- Perrett, D.I., Rolls, E.T. & Caan, W. (1982) Visual neurons responsive to faces in the monkey temporal cortex. *Exp. Brain Res.*, **47**, 329–342.
- Polyak, S. (1941) *The Retina*. University of Chicago Press, Chicago.
- Rolls, E.T. (1992) Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Phil. Trans. R. Soc. Lond. B*, **335**, 11–21.
- Rolls, E.T. (2000) Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron*, **27**, 205–218.
- Rolls, E.T. (2005) *Emotion Explained*. Oxford University Press, Oxford.
- Rolls, E.T., Aggelopoulos, N.C. & Zheng, F. (2003) The receptive fields of inferior temporal cortex neurons in natural scenes. *J. Neurosci.*, **23**, 339–348.
- Rolls, E.T. & Baylis, G.C. (1986) Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Exp. Brain Res.*, **65**, 38–48.
- Rolls, E.T., Baylis, G.C. & Leonard, C.M. (1985) Role of low and high spatial frequencies in the face-selective responses of neurons in the cortex in the superior temporal sulcus in the monkey. *Vis. Res.*, **25**, 1021–1035.
- Rolls, E.T. & Cowey, A. (1970) Topography of the retina and striate cortex and its relationship to visual acuity in rhesus monkeys and squirrel monkeys. *Exp. Brain Res.*, **10**, 298–310.
- Rolls, E.T. & Deco, G. (2002) *Computational Neuroscience of Vision*. Oxford University Press, Oxford.
- Rolls, E.T., Sanghera, M.K. & Roper-Hall, A. (1979) The latency of activation of neurons in the lateral hypothalamus and substantia innominata during feeding in the monkey. *Brain Res.*, **164**, 121–135.
- Rolls, E.T. & Tovee, M.J. (1995a) The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the visual field. *Exp. Brain Res.*, **103**, 409–420.
- Rolls, E.T. & Tovee, M.J. (1995b) Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J. Neurophysiol.*, **73**, 713–726.
- Rolls, E.T., Yaxley, S. & Sienkiewicz, Z.J. (1990) Gustatory responses of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *J. Neurophysiol.*, **64**, 1055–1066.
- Sato, T. (1989) Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Exp. Brain Res.*, **77**, 23–30.
- Sato, T. (1995) Interactions between two different visual stimuli in the receptive fields of inferior temporal neurons in macaques during matching behaviours. *Exp. Brain Res.*, **105**, 209–219.
- Sheinberg, D.L. & Logothetis, N.K. (2001) Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. *J. Neurosci.*, **21**, 1340–1350.
- Siegel, S. (1956) *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York.
- Singer, W. (1999) Neuronal synchrony: a versatile code for the definition of relations? *Neuron*, **24**, 49–65.
- Tanaka, K. (1996) Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.*, **19**, 109–139.
- Tovee, M.J., Rolls, E.T. & Azzopardi, P. (1994) Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. *J. Neurophysiol.*, **72**, 1049–1060.
- Tovee, M.J., Rolls, E.T., Treves, A. & Bellis, R.P. (1993) Information encoding and the responses of single neurons in the primate temporal visual cortex. *J. Neurophysiol.*, **70**, 640–654.
- Trappenberg, T.P., Rolls, E.T. & Stringer, S.M. (2002) Effective size of receptive fields of inferior temporal cortex neurons in natural scenes. In Dietterich, T.G., Becker, S. & Ghahramani, Z. (Eds), *Advances in Neural Information Processing Systems 14*. MIT Press, Cambridge, MA, pp. 293–300.
- Treves, A., Panzeri, S., Rolls, E.T., Booth, M. & Waksman, E.A. (1999) Firing rate distributions and efficiency of information transmission of inferior temporal cortex neurons to natural visual stimuli. *Neural Comput.*, **11**, 611–641.
- Vogels, R. (1999) Categorization of complex visual images by rhesus monkeys. Part 2: single-cell study. *Eur. J. Neurosci.*, **11**, 1239–1255.