JOURNAL OF NEUROPHYSIOLOGY
Vol. 71, No. 3, March 1994. *Printed in U.S.A.*

Neuronal Selectivities to Complex Object Features in the Ventral Visual Pathway of the Macaque Cerebral Cortex

EUCALY KOBATAKE AND KEIJI TANAKA
Laboratory for Neural Information Processing, Frontier Research Program, and Information Science Laboratory, The Institute of Physical and Chemical Research, Wako-shi, Saitama 351-01, Japan

SUMMARY AND CONCLUSIONS

1. To infer relative roles of cortical areas at different stages of the ventral visual pathway, we quantitatively examined visual responses of cells in V2, V4, the posterior part of the inferotemporal cortex (posterior IT), and the anterior part of the inferotemporal cortex (anterior IT), using anesthetized macaque monkeys.

2. The critical feature for the activation was first determined for each recorded cell by using a reduction method. We started from images of three-dimensional complex objects and simplified the image of effective stimuli step by step by eliminating a part of the features present in the image. The simplest feature that maximally activated the cell was determined as the critical feature. The response to the critical feature was then compared with responses of the same cell to a routine set of 32 simple stimuli, which included white and black bars of four different orientations and squares or spots of four different colors.

3. Cells that responded maximally to particular complex object features were found in posterior IT and V4 as well as in anterior IT. The cells in posterior IT and V4 were, however, different from the cells in anterior IT in that many of them responded to some extent to some simple features, that the size of the receptive field was small, and that they intermingled in single penetrations with cells that responded maximally to some simple features. The complex critical features in posterior IT and V4 varied; they consisted of complex shapes, combinations of a shape and texture, and combinations of a shape and color.

4. We suggest that local neuronal networks in V4 and posterior IT play an essential role in the formation of selective responses to complex object features.

INTRODUCTION

Among the multiple visual cortical areas of the macaque monkey, V1, V2, V4, the posterior part of the inferotemporal cortex (posterior IT or TEO), and the anterior part of the inferotemporal cortex (anterior IT or TE) form a serial pathway, which is called the "ventral visual cortical pathway." This ventral pathway is thought to be responsible for "object vision," i.e., discrimination and recognition of visual images of objects, because a bilateral lesion of the inferotemporal cortex resulted in a severe deficit in learning tasks that required these functions (for review see Dean 1976; Gross 1972; Mishkin 1982). The indispensable role of the ventral pathway for the object vision has been further supported by findings of selective cell responses to complex object features (Desimone et al. 1984; Gross et al. 1969, 1972; Schwartz et al. 1983; Tanaka et al. 1991) and of the columnar organization in anterior IT (Fujita et al. 1992). However, little is known of individual specific roles of the areas at different stages of the pathway. Theoretical studies (e.g., Marr 1982) propose that the visual recognition of objects requires many steps of computation; thus we wish to assign different kinds of computation to different stages of the pathway. Although a functional dissociation between posterior IT and anterior IT was suggested on the basis of lesion studies (Iwai and Mishkin 1968), this concept has not been further developed, probably because of the lack of related data at the cellular level. As a step towards this goal, we compared selectivity of cell responses for complex object features among the cortical areas in the ventral visual cortical pathway.

It has been well documented that cells in anterior IT respond selectively to particular complex object features. We asked whether cells in the earlier stages also respond selectively to complex object features, and if so, how they are different from the cells in anterior IT. We applied the same reduction approach to study the selectivity of cells in V2, V4, and posterior IT, as well as anterior IT. Effective stimuli among complex object stimuli were first determined, and then the complexity of the effective stimuli was reduced to identify the feature critical for the activation. We then quantified the distinctiveness of the selectivity to complex object features; a set of simple stimuli was prepared and the response to the critical feature was compared with responses of the same cell to the set stimuli. Although the comparison of the critical features between anterior IT and some of the former areas had already been carried out in a previous study (Tanaka et al. 1991), the distinctiveness of the selectivity was quantified for the first time in the present study. Some of the results were previously reported in abstract form (Kobatake and Tanaka 1991, 1992).

METHODS

Preparation and recording

The general methods of preparation and recording were similar to those described previously (Tanaka et al. 1991). Two adult Japanese monkeys (*Macaca fuscata; monkey 1* and *monkey 2*), each weighing 5.5 kg, were prepared for repeated recordings. The initial aseptic surgery was performed under anesthesia with pentobarbital sodium. A brass block for head fixation was attached to the top of the skull with stainless steel bolts and acrylic resin, two stainless steel bolts for electroencephalographic recording were implanted in the skull, and the right lateral surface of the skull was exposed and covered with resin for later unit recordings. An antibiotic was given daily for 1 wk after the surgery.

Before the first session of unit recordings the optics of the eyes were measured to select appropriate contact lenses. The curvature of the cornea was measured and after a contact lens with appro-

0022-3077/94 $3.00 Copyright © 1994 The American Physiological Society

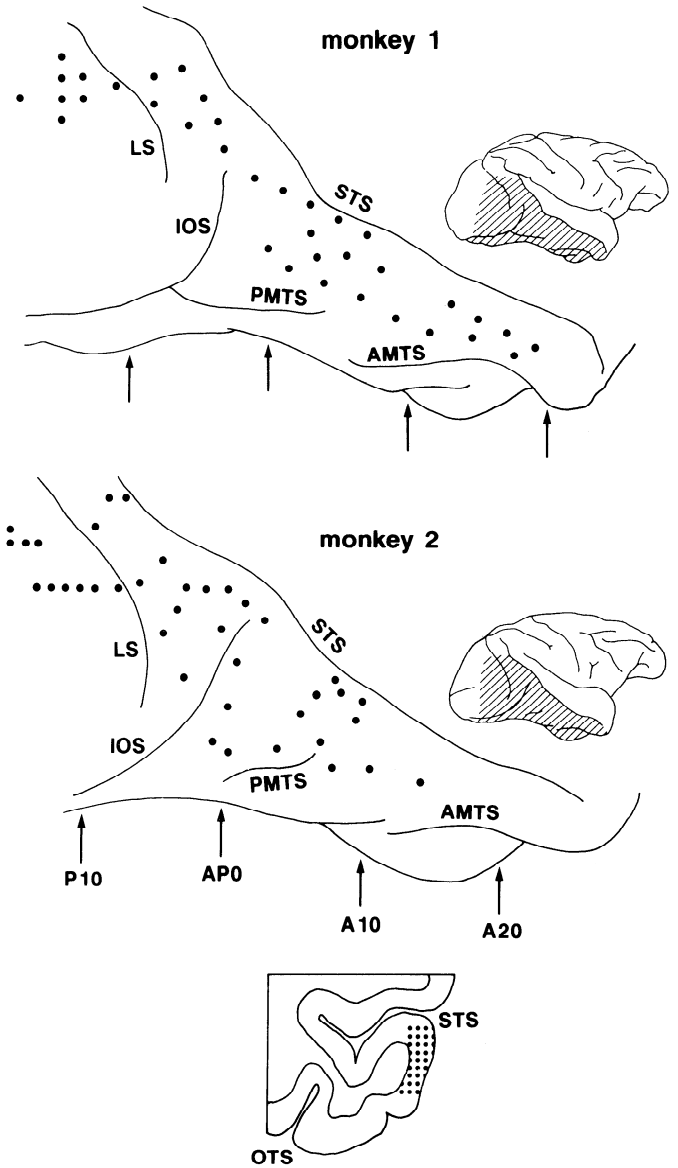

FIG. 1. Distribution of penetrations. *Top two*: positions of successful electrode penetrations are plotted at the entrance on the lateral view of the cortex. The locations of the depicted regions in the brain are indicated by the hatched regions in the *insets*. In these penetrations quantitative data were obtained from ≥1 cell. Recordings in the penetrations posterior to the lip of the lunate sulcus were performed in the lateral bank of the sulcus. P10, AP0, A10, and A20, level of posterior 10, 0, anterior 10, and anterior 20 in Horsley-Clarke coordinates. *Bottom*: coronal section at A10 of *monkey 2*. Dotted region: region covered by the penetrations made at this and nearby levels. Coronal sections are not available for the brain of *monkey 1*, which was cut in horizontal planes. STS, superior temporal sulcus; LS, lunate sulcus; IOS, inferior occipital sulcus; PMTS, posterior middle temporal sulcus; AMTS, anterior middle temporal sulcus; OTS, occipitotemporal sulcus.

priate curvature was placed on the cornea the reflectance was measured to determine the power of the lens with which images at a distance of 57 cm from the cornea were focused on the retina. Photographs of the fundus were taken to determine the position of the fovea.

Recordings were made once a week on each monkey. The recording session began with the induction of anesthesia with ketamine hydrochloride (10 mg/kg im). An endotracheal cannula was inserted through the tracheal opening and a small hole (7–10 mm diam) was drilled in the resin-coated skull under the anesthesia with isoflurane. Throughout the recording the animal was immobilized with pancuronium bromide and anesthesia was maintained by artificial ventilation with a gas mixture of N_2O-O_2 and isoflurane. To reduce salivation, 0.5 mg atropine sulfate was injected subcutaneously every 3 h. Extracellular single-cell recordings were made using a glass-coated Elgiloy electrode (2–3.5 MΩ at 1 kHz). The electrode was inserted into the brain through a pinhole in the dura and advanced in the frontal plane medially (in *monkey 2*) or medioventrally at an angle of 5° from the horizontal plane (in *monkey 1*). The pinhole was made by a needle of stainless steel with a shaft 0.7 mm diam. The exposed dura was covered with paraffin to prevent it from drying and to reduce movements of the brain caused by pulsation or respiration. The position of each penetration was determined with reference to a point marked on the resin-coated skull. The hole in the skull was filled with resin after the recording was finished. Within a few hours after the last injection of the muscle relaxant, spontaneous respiration recovered to normal. The monkey was returned to the home cage after an injection of antibiotic. The length of time from the induction of anesthesia to the recovery of spontaneous respiration was <16 h.

Only one or two penetrations were made in one recording session, but mapping in the same hemisphere was continued for 9 mo (*monkey 1*) and 10 mo (*monkey 2*). Figure 1 shows the distribution of penetrations in the two hemispheres. All the penetrations were made in the right hemisphere. From anterior to posterior, the mapped regions almost completely covered the lateral surface of the inferotemporal cortex, the prelunate gyrus, which is known to

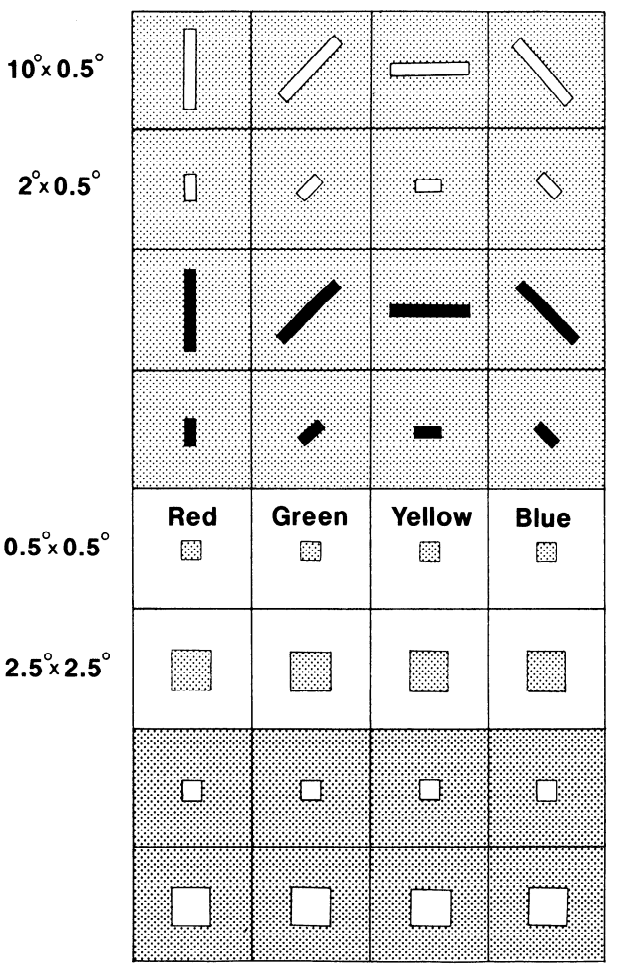

FIG. 2. Set of simple stimuli. To measure the distinctiveness of the selectivity, the response to the complex critical feature was compared with responses of the same cell to this set of simple stimuli.

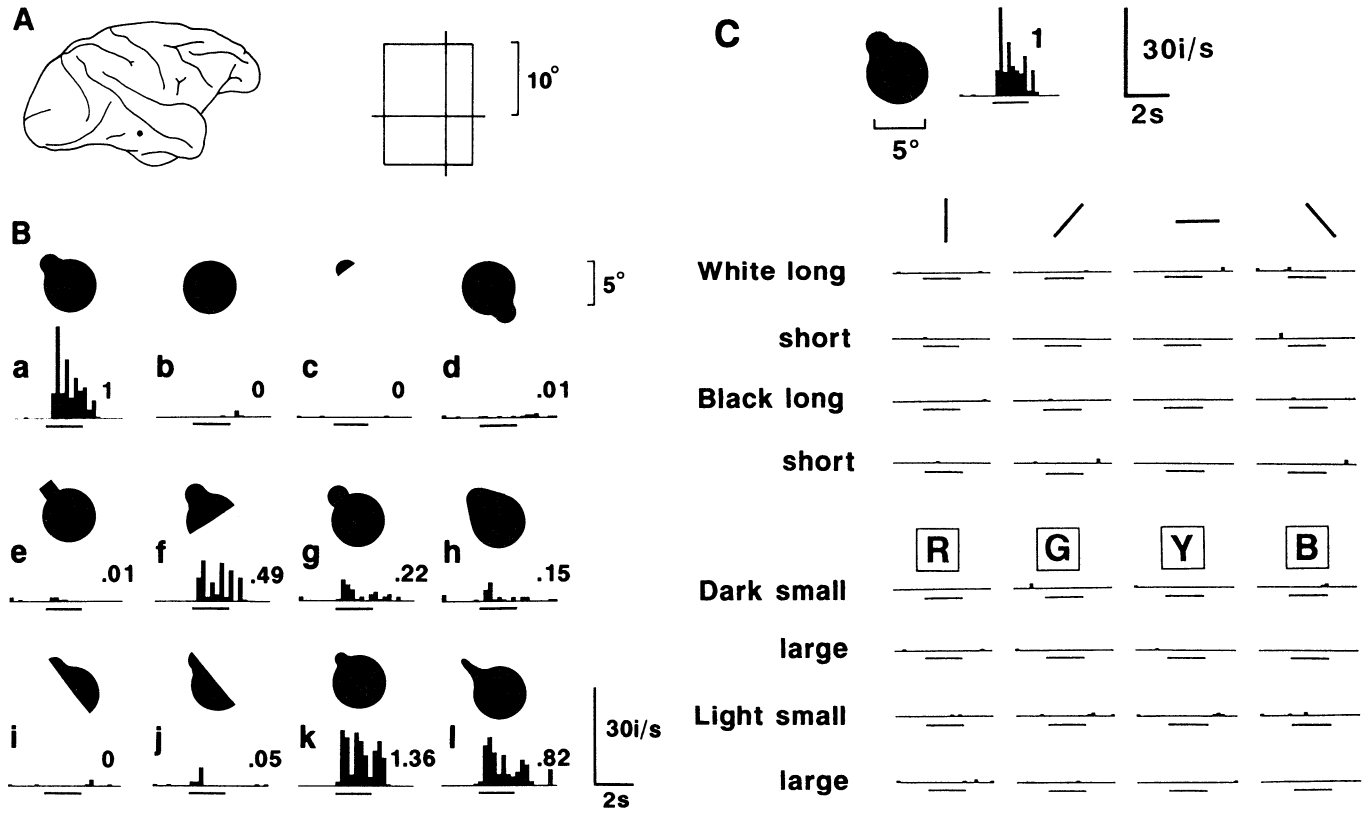

FIG. 3. Example of a cell that responded to a complex critical feature with distinctive selectivity. *A*: location of the penetration and the receptive field. *B*: responses to the critical feature and related stimuli. Responses were averaged over 10 stimulus presentations in this and the following figures. Underlines: time of stimulus presentation. Values above individual peristimulus time histograms (PSTHs): magnitude of the response normalized by that of the response in *a*. The cell responded to a stimulus configuration in which a rounded head was pointing to the upper left direction from a rounded body with a smooth concave neck (*a*). The body or the head by itself did not evoke any responses (*b* and *c*). The direction of projection was critical (*d*). The head had to be rounded because the response disappeared when the rounded head was replaced by a square (*e*). The body had to be rounded because the response decreased by 51% when the body was cut in half (*f*). The neck had to be smooth and concave because the response decreased by 78 or 85% when the neck was replaced by that with sharp corners (*g*) or the neck was straight (*h*). The critical feature was neither the right upper contour alone nor the left lower contour alone because either half of the stimulus did not evoke responses (*i* and *j*). The width and length of the projection were not critical (*k* and *l*). The responses in *f*–*h* were significantly weaker than the response in *a* ($P < 0.01$, Kolmogorov-Smirnov test). *C*: responses to the set of simple stimuli. They are arranged as in Fig. 2. R, red; G, green; Y, yellow; B, blue.

be occupied by V4, and the lateral bank of the lunate sulcus, which is occupied by V2. The mapping was started at the anterior end and was gradually moved to the posterior.

After the last recording session several needles were left in the brain and the monkey was deeply anesthetized with pentobarbital sodium and then perfused intracardially. Frozen sections were cut 50 μm thick in the frontal plane and a series of sections was taken at 250-μm intervals to be stained with cresyl violet.

Visual stimuli and procedure for individual cells

The pupils were dilated and the lenses relaxed by local application of 0.5% tropicamide-0.5% phenylephrine. The corneas were covered with contact lenses of appropriate power with artificial pupils 3 mm diam. A television display (CMM20-11, Shibasoku) was placed at a distance of 57 cm from the corneas. Several retinal landmarks, such as the intersection of vessels and the center of the optic disk, were projected onto the display by the use of a reversible retinoscope, and the position of the fovea was determined geometrically by referring to the photographs of the fundus. The monkey saw the stimuli monocularly, usually by the eye contralateral to the recording side. The other eye was occluded by a removable opaque plate. The picture on the television display was 31 × 29 cm in full size, composed of 60 fields per second with interlacing, 512 × 480 pixels, and 255 levels for red, green, and blue. The brightest white and the darkest black (with the room lights on) were 76 and 0.96 cd/m^2, respectively. All the experiments were conducted with the room lights (fluorescent lamps) on.

Stimulus selectivity of individual cells was examined in three steps. Responses of a recorded cell were first tested with a routine set of stimuli, which were presented by hand in front of the display, and the effective stimuli were listed. The routine set was composed of *1*) bars and disks of various sizes and colors; *2*) regular geometric patterns, such as stripes, dot patterns, concentric rings, and patterns such as windmills, drawn on paper; *3*) plastic and sponge spheres, sponge cubes, pasteboard cylinders, and feather brooms of various colors; *4*) three-dimensional animal stimuli made of vinyl, cloth, or plastic, including a tiger, tabby cat, spotted dog, zebra, giraffe, gorilla, hawk, duck, frog, raccoon, monkey, human head, and human hand; *5*) three-dimensional plant stimuli made of plastic, including banana, apple, ear of corn, pineapple, grapefruit, melon, cabbage, carrot, potato, cucumber, watermelon, eggplant, onion, pear, potato, bunch of grapes, ivy, cut piece of apple, pepper, and pumpkin; and *6*) the experimenter's hand, body, and face. Various sides of the objects were presented with various orientations.

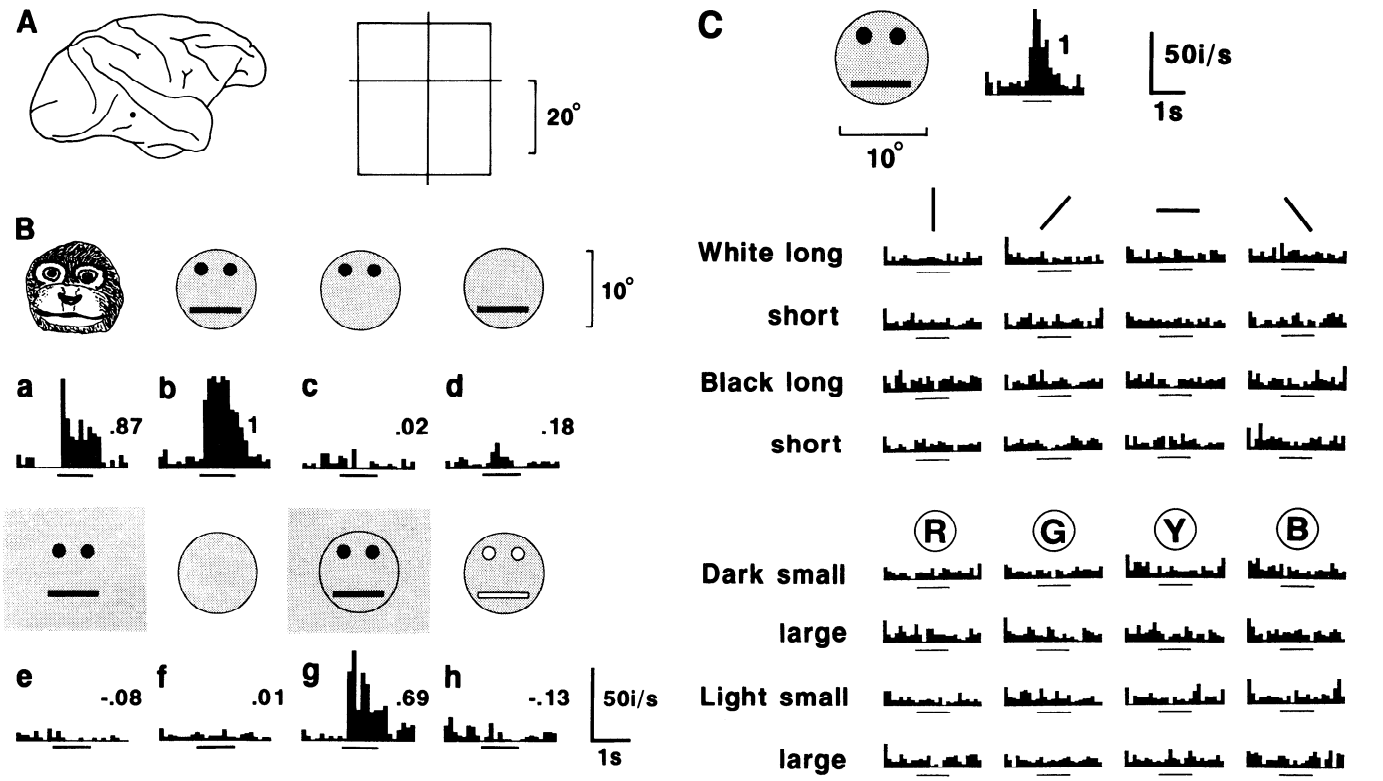

FIG. 4. Second example of a cell that responded to a complex critical feature with distinctive selectivity. The cell responded strongly to the face of a monkey toy (*a*) and the critical feature was determined as a configuration in which 2 black spots and 1 horizontal black bar were arranged in a gray disk (*b*). Both the bar and the spots were indispensable (*c* and *d*) and the circular outline was essential (*e*). The contrast between the inside and outside of the circular contour was not critical (*g*), although the response in *g* was significantly weaker than that in *b* ($P < 0.05$). The spots and bar had to be darker than the background within the outline (*h*). The small response in *d* was significantly weaker than that in *b* ($P < 0.01$).

Even for a cell that reliably responded to some simple stimuli, such as bars, disks, and gratings, we searched for more effective stimuli among the more complex stimuli. In particular, stimuli that contained the effective simple feature were carefully examined. By doing this, we succeeded in finding a unique class of cells that showed maximal responses to some complex features and weaker but definite responses to some simple stimuli.

The images of the effective objects were then presented on the television display using a CCD video camera (DXC-325, Sony) to examine whether the pictures on the television display activated the cells. Some cells in the inferotemporal cortex hardly responded to television pictures, although presentation of the actual object seemed to evoke reliable responses. Possible reasons for this attenuated response to television pictures were the limitations in the spatial resolution, the limitations in the dynamic range of brightness, and the subtle differences in color. If a cell responded reliably to the television picture we recorded the picture and stored it in the memory of an image processing computer (Nexus 600-M4).

We then determined which component or combination of components in the image was essential for the activation. The picture was simplified by step-by-step removal of some features. If a cell responded to a simplified image as strongly as to the original image, we assumed that the excluded features were not necessary for the activation. This procedure was first performed qualitatively by monitoring the responses using an audio monitor, but the responses were later evaluated quantitatively by making peristimulus time histograms. If a cell showed comparably strong responses to more than two different pictures we examined features common to the pictures and compared the effectiveness of the picture with that of the original pictures. Finally, the simplest or most elemental picture for the maximal activation of the cell was determined. We call it the "critical feature" for the activation of the cell. Cells often showed stronger responses to simplified images than to the original images (see Fig. 6 of Tanaka et al. 1991), possibly because of an increase in the contrast of features or the elimination of inhibitory effects of unrelated features.

Finally, to evaluate the distinctiveness of the selectivity, the response to the critical feature was quantitatively compared with responses of the same cell to a routine set of simple stimuli. The set, which is shown in Fig. 2, was composed of vertical bars, horizontal bars, and bars at 45° angles, and squares or spots of four different colors. Two sizes of each stimulus were included, as well as stimuli with darker and lighter backgrounds. The luminance of the white bars, that of the black bars, and that of the gray background were 52.5, 1.4, and 6.1 cd/m^2. The luminance of color stimuli and that of their background were 2.5 and 17.2 cd/m^2 for the light background stimuli and 5.2 and 1.4 cd/m^2 for the dark background stimuli. We added a few simple stimuli whose parameters were adjusted to those of the critical stimulus if the cell required precise adjustment of the size, orientation, or color of the critical stimulus. A few simple textures, i.e., straight stripes and dot patterns, were added if the critical stimulus seemed to include some texture. These simple textures were included as simple features in this paper, because cells that responded selectively to straight stripes have previously been found in V1 and V2 (Albrecht et al. 1980; von der Heydt et al. 1992). For cells whose critical feature was identified as a simple feature we selected a few of the most effective object stimuli, obtained their images, and combined these pictures with the set stimuli.

There is an important ramification of the use of the set of simple stimuli. In some cells the quantitative run with the set stimuli revealed some less distinct selectivity of the cell and we obtained a

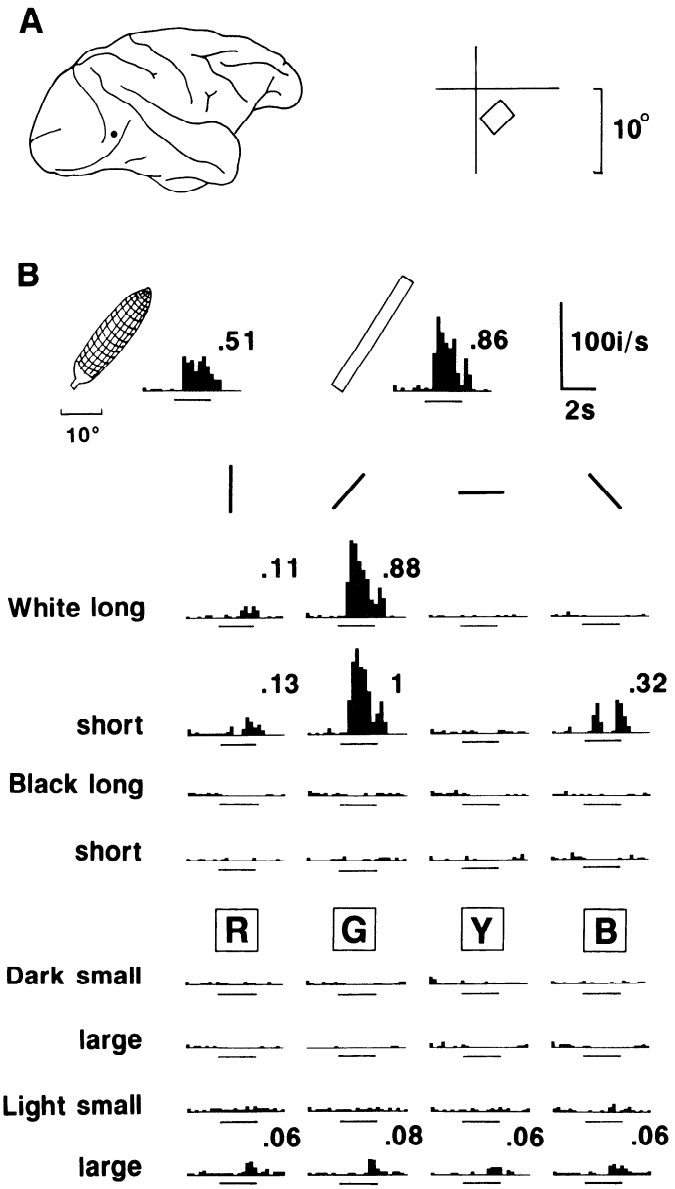

FIG. 5. Example of a cell that responded maximally to a simple feature. In this and the following figures (Figs. 6, 7, and 11), the stimuli that are brighter than the background are drawn by outlines only. The small responses to the vertical bars, the left-up oblique short bar, and colored squares were all significantly weaker than any of the responses to the right-up oblique white bars.

more effective stimulus by adding the newly found feature to the stimulus previously assumed to elicit the optimal response. This feedback from the quantitative run sometimes resulted in a change in the critical feature from simple to complex. This indicates that the reduction approach to the determination of the selectivity of cells was not perfect.

The extent of the receptive field was determined by use of the optimal stimulus. We plotted the border by the center of the stimulus. The receptive fields would have been smaller if the inner edge of the stimulus had been used as for "the minimum response field" (Barlow et al. 1967) and they would have been larger if the outer edge of the stimulus had been used. To decrease the differences we used the smallest stimulus to which the cell responded maximally. The positions on the tangent screen were converted into those in visual angles and the area of the receptive field was calculated as if it were elliptical in shape.

Quantitative evaluation of responses

For both the procedure to determine the critical feature and that to evaluate the distinctiveness of selectivity, different pictures to compare were always intermixed and presented in a cyclic order 10 times. They were presented with cyclic translation (without a change of the orientation) for 1.1 s with 2.2-s intervals or for 2.2 s with 4.4-s intervals. The radius of the cyclic translation was 0.58 or 0.29° and the period of the movement was 0.86 or 0.43 s. The stimuli were intermixed to compensate for the slow change in excitability and to prevent habituation to the effective stimuli. The magnitude of the response was represented by the averaged firing rate during the stimulus presentation minus the averaged firing rate for 1 s immediately before the stimulus presentation. The window for the response was shifted within the range of 50–250 ms from the exact time of stimulus presentation so that the maximum firing rate was taken. Differences between two responses were analyzed by comparing two sets of 10 individual responses with Kolmogorov-Smirnov test and they were regarded to be different if the difference was significant with $P < 0.05$. The significance of a response itself was examined by comparing the set of 10 averaged firing rates during stimulus presentation with the set of 10 averaged firing rates immediately before the stimulus presentation.

RESULTS

The database of this study is composed of 228 cells for which the quantitative examination of the distinctiveness of selectivity was completed. Their responses to the individual optimal stimuli were significantly stronger than their spontaneous discharges ($P < 0.01$). The receptive fields of 30 additional cells were also included in the data of the receptive field. These latter cells were clearly responsive, but the recording time was too short to allow quantitative data to be obtained regarding the distinctiveness of selectivity.

Cells that responded maximally to complex object features were found in V4 and posterior IT as well as in anterior IT. However, the distinctiveness of the selectivity differed between the cells in V4 and posterior IT and the cells in anterior IT.

Cells in anterior IT showed distinctive selectivity to particular complex features. The cell shown in Fig. 3 responded maximally to a pear model within the routine set of three-dimensional object stimuli, and the critical feature for the activation was determined as "a rounded projection from a rounded body with a concave smooth neck" after the reduction process. The receptive field was large ($12.3 \times 16.3°$) and included the fovea. When the set of simple stimuli was presented in combination with the critical feature, no responses were evoked by any of the simple stimuli. Although this first cell yielded very few spontaneous discharges, the absence of responses to simple stimuli was not necessarily accompanied by low spontaneous activity in other cells.

Another anterior IT cell (Fig. 4), which showed rather frequent spontaneous discharges, also showed distinctive selectivity. This second cell responded to the face of a toy monkey. After the reduction process the critical feature was determined as an arrangement of two black spots and one black horizontal bar in a gray disk. The presentation of the simple stimuli did not elicit an increase in the firing rate. These two cells (Figs. 3 and 4) represent a large majority

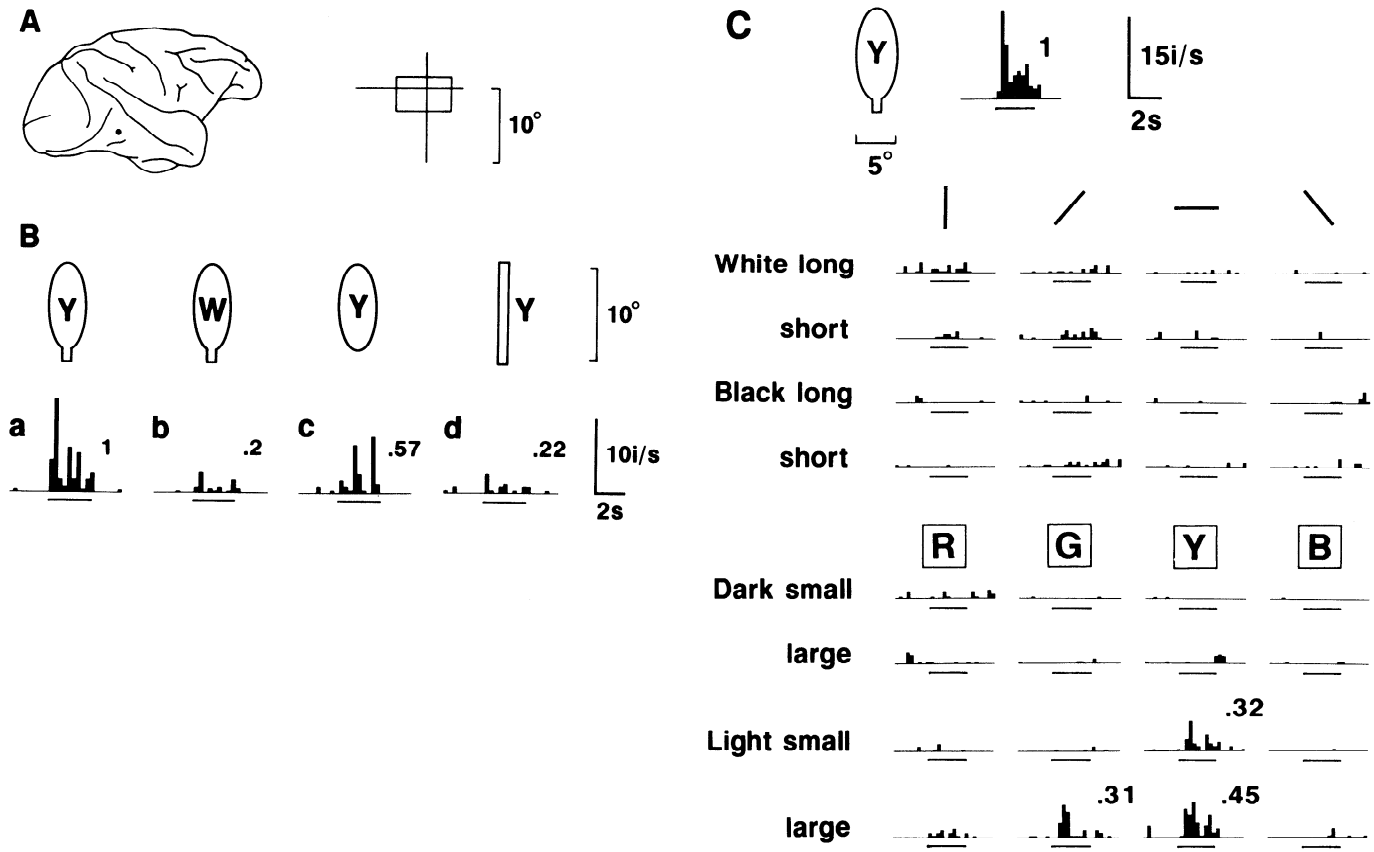

FIG. 6. Example of a cell that showed maximal response to a complex critical feature and weaker but definite responses to some of the simple stimuli. *B*: critical feature was determined as a yellow ellipse with a small downward projection (*a*). A white stimulus of the same shape, a yellow ellipse, and a yellow vertical bar evoked some response (*b–d*), but these responses were significantly weaker than the response in *a* (all $P < 0.01$). *C*: there were responses to yellow squares that were brighter than the background and the large green square that was brighter than the background, but they were significantly weaker than the response to the critical feature (all $P < 0.05$). W, white.

(~3/4) of the cells in anterior IT. They responded maximally to a particular complex feature and showed no or very weak (<1/4 of the maximum) responses to any simple stimuli.

When the position of penetrations moved to the posterior, the proportion of cells with distinctive selectivity to complex features decreased and the size of the receptive field decreased. This change will be described in detail later in Figs. 8 and 9. Because the transition was more abrupt than gradual, we used this physiological change to discriminate the anterior and posterior parts of the inferotemporal cortex. We denote the part of the inferotemporal cortex posterior to this transition "posterior IT" and the anterior part "anterior IT" (Tanaka et al. 1991). Posterior IT may correspond to the cytoarchitectural area TEO and anterior IT to TE. We did not further divide anterior IT, and so our anterior IT may include both TE posterior (TEp, CIT) and TE anterior (TEa, AIT).

About one half of the cells in posterior IT and V4 could be maximally activated by some simple stimuli. There were effective object stimuli, but some simple features common to them elicited a response stronger than or as strong as that of the object stimuli. These cells are exemplified by a cell that was recorded in V4 (Fig. 5). It responded to many elongated objects whose long axes were oriented obliquely (lower left to upper right). An ear of corn was used for the quantitative test. When the set of simple stimuli was shown to the monkey, white bars, both long and short, with the same orientation as that of the effective object stimuli evoked responses greater than the response to the ear of corn. A larger white bar whose size and orientation were adjusted to those of the ear of corn was just as effective as the white bars in the set. White bars of other orientations, black bars, and colored squares were not effective. Thus the selectivity of the cell can be fully explained within the domains of orientation and contrast polarity. The elongated objects were effective because the images contained a component that fulfilled these conditions. Although it was not the case in this particular cell, complementary simple stimuli (see METHODS) sometimes evoked a response stronger than those to the set stimuli. A straight stripe or a dot pattern, which was added as a complementary simple stimulus, evoked the strongest response in 11 cells (stripe in 9 cells and dot pattern in 2 cells). These simple textures were included as simple features in this study because cells that responded selectively to straight stripes have previously been found in V1 and V2 (Albrecht et al. 1980; von der Heydt et al. 1992).

The other half of the cells in V4 and posterior IT were maximally activated only by stimuli including some complex critical features. Some of these cells did not respond to any simple stimuli at all, as the cells in anterior IT, whereas other cells showed moderately strong responses to some simple stimuli. Two examples of the latter cells are discussed,

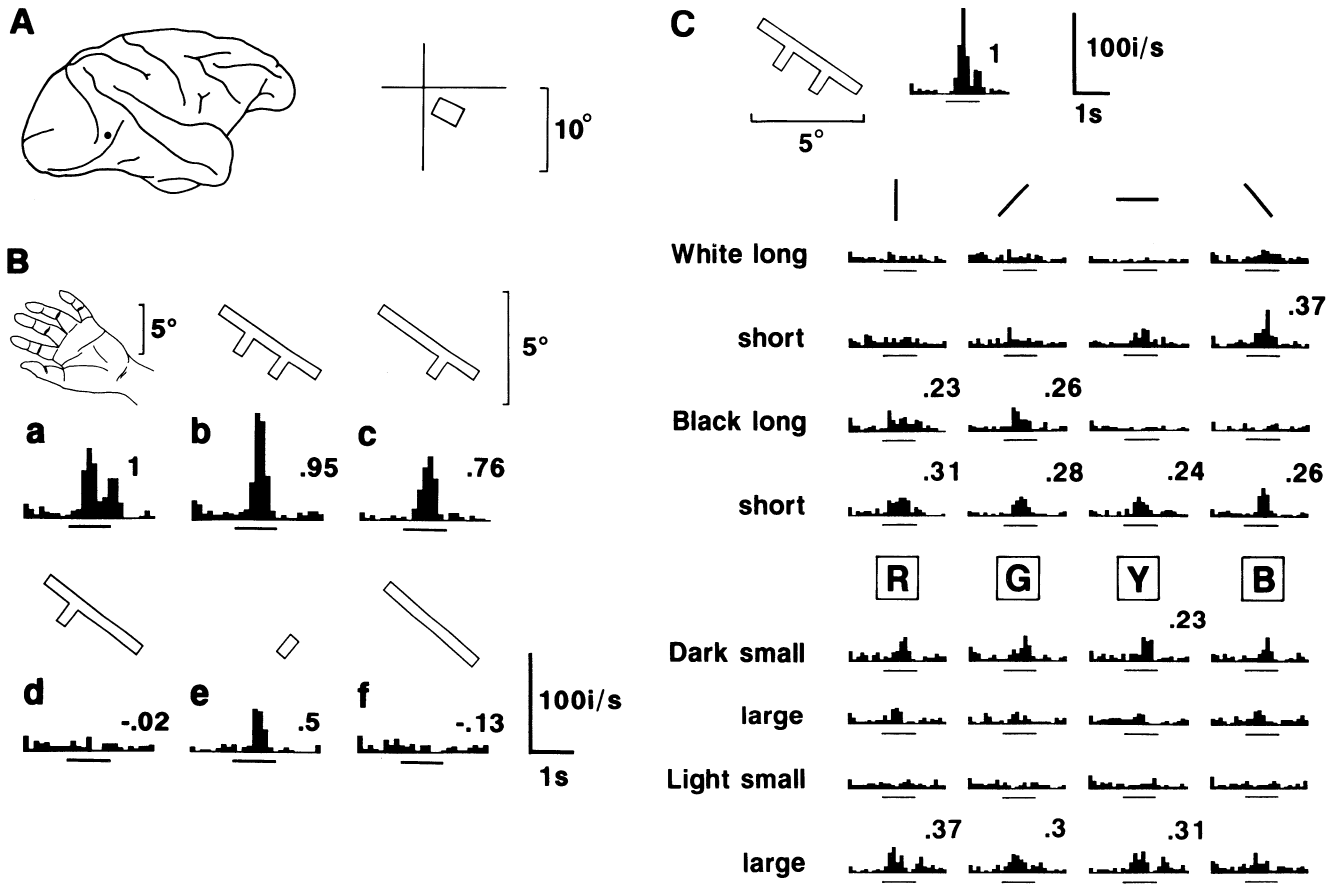

FIG. 7. Second example of a cell that showed maximal response to a complex critical feature and weaker but definite responses to some of the simple stimuli. *B*: cell responded strongly to the palm of the hand (*a*) and the critical feature was determined as a combination of 2 short bars and 1 long bar (*b*). A part of the combination evoked a weaker or no response (*c–f*). The responses in *c* and *e* were significantly weaker than the response in *b* (both $P < 0.01$). *C*: there were responses to many of the simple stimuli, but they were all significantly weaker than the response to the critical feature (all $P < 0.01$).

one recorded from posterior IT (Fig. 6) and the other from V4 (Fig. 7).

The cell shown in Fig. 6 was recorded from a penetration made close to the posterior middle temporal sulcus. The critical feature was a yellow ellipse with a small downward protrusion, which was reduced from a plastic model of an ear of corn. When the set of simple stimuli was presented, both small and large yellow squares brighter than the background evoked 32–45% of the response to the critical feature. The receptive field was smaller ($7.7 \times 4.7°$) than those of the anterior IT cells.

The cell shown in Fig. 7 was recorded from a penetration made in the ventral part of the prelunate gyrus, which is known to be occupied by V4. The cell responded strongly to the hand of a mannequin, and the critical feature was determined as a combination of three bars, which may correspond to the upper edge and lines at the joints of a finger. When the set of simple stimuli was presented, moderately strong responses ($\leq$37% of the response to the critical feature) were evoked by many of the stimuli.

Thus cells in posterior IT and V4 showed great variation in selectivity. Some cells responded only to stimuli that included some particular complex features. Other cells responded maximally to some complex features but also showed moderately strong responses to some simple stimuli. The remaining cells did not require complex features for their maximal activation, that is, they were maximally activated by some simple features. These diverse cells, however, did not make distinctive groups. They rather made a continuum in the distinctiveness of selectivity to complex features.

To objectively examine the distribution of cells with varying selectivity for complex features we quantified the distinctiveness of the selectivity by the ratio of the maximum responses to simple stimuli to the maximum response to all the stimuli combined in the run (S_{max}/MAX). The ratio was 1 for a cell that maximally responded to a simple stimulus and it was close to 0 for a cell that exclusively responded to a particular complex feature. In Fig. 8, cells are arbitrarily divided at 0.25, 0.5, and 0.75 into four ranges and cells in different ranges are represented by different symbols. Cells with a ratio <0.25 predominated in the anterior part of the inferotemporal cortex, cells in the four ranges intermingled in the posterior part of the inferotemporal cortex and the prelunate gyrus (V4), and cells with a ratio >0.75 predominated in the lateral bank of the lunate sulcus (V2).

Cells with S_{max}/MAX ratios in different ranges intermingled in single penetrations in the posterior-inferotemporal cortex and V4. This can be seen in Fig. 8, in which cells recorded in single penetrations are grouped together. Of 35 penetrations in which more than two cells were studied

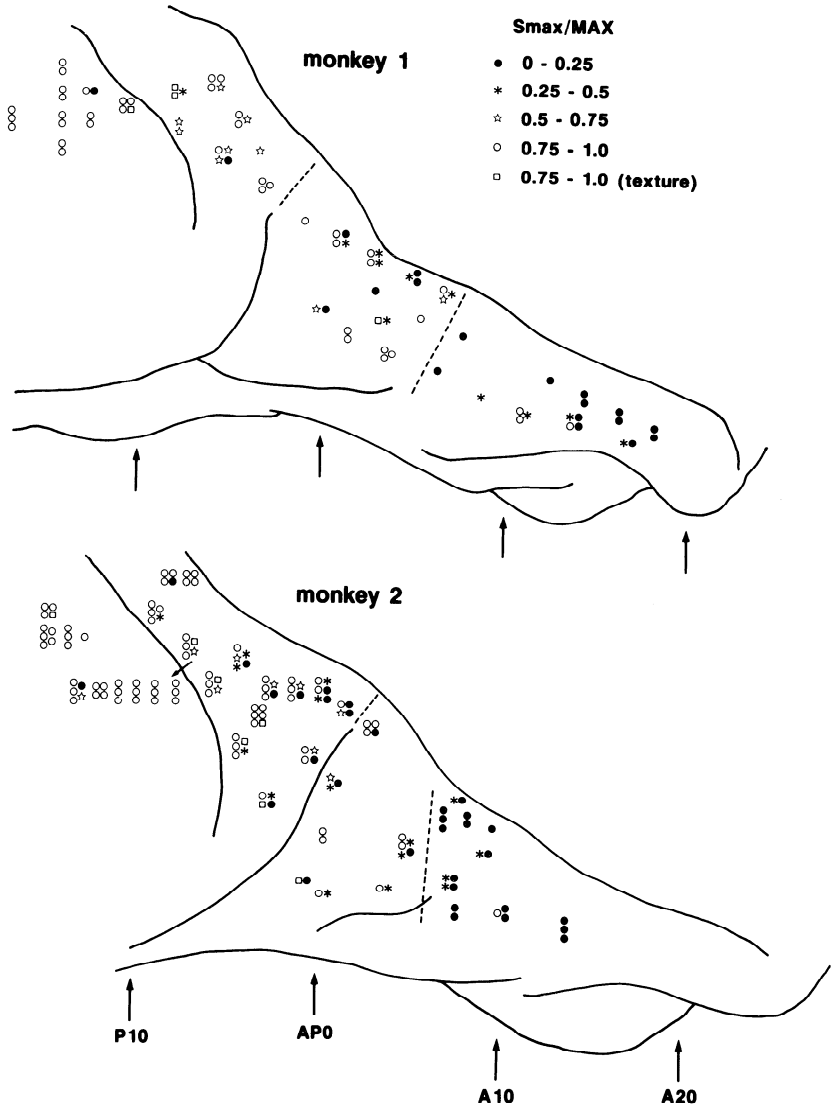

FIG.. 8. Distribution of cells with varying selectivity to complex features. To quantify the distinctiveness of the selectivity to complex features, we calculated the ratio of the maximum of the responses to simple stimuli to the maximum of the responses to all the stimuli (S_{max}/MAX). A smaller ratio represents more distinctive selectivity. Cells are indicated by 4 different symbols according to the ratio as shown at the *top right corner.* Among the cells with a ratio >0.75, those that responded to some simple textures but not to bars or spots are indicated by open squares. Broken lines: the border between anterior part of inferotemporal cortex (anterior IT) and posterior part of inferotemporal cortex (posterior IT) and that between posterior IT and V4. Cells recorded from a single penetration are grouped together. The 5 V4 cells recorded from the medial bank of the lunate sulcus (in monkey 2) are shifted to the prelunate gyrus (indicated by an arrow). The V2 cells were recorded from the lateral bank of the lunate sulcus. The depicted regions are the same as those of Fig. 1.

quantitatively, 23 penetrations included both cells with a ratio >0.5 and cells with a ratio <0.5.

The change in the balance between cells in different ranges from the anterior part to the posterior part of the inferotemporal cortex was accompanied by a change in the size of the receptive field. In Fig. 9 the receptive fields of cells recorded in single penetrations are superimposed. The receptive fields in the anterior inferotemporal cortex were large and included the fovea, but the receptive fields in the posterior inferotemporal cortex were much smaller and they did not necessarily include the fovea. This transition was abrupt in *monkey 1* but appeared to be more gradual in *monkey 2.* The difference in the size of the receptive fields between the anterior and posterior parts of the inferotemporal cortex has previously been reported (Boussaoud et al. 1991; Desimone and Gross 1979; Tanaka et al. 1991). The receptive fields in the prelunate gyrus were as small as those in the posterior inferotemporal cortex. As reported previously (Gattass et al. 1988; Maguire and Baizer 1984), there was a gradual shift from the fovea to the periphery as the position of penetration moved from the ventral to the dorsal region in the prelunate gyrus. The receptive fields in V2 were the smallest.

On the basis of the transitions in the selectivity and in the size of the receptive fields we drew borders between anterior IT and posterior IT as shown in Figs. 8 and 9 by the broken lines. The borders between V4 and posterior IT were tentatively drawn by extrapolating the anterior tip of the inferior occipital sulcus. Because we did not record cells in the region ventral to the inferior occipital sulcus, this border between V4 and posterior IT may mostly correspond to the border between V4 and TEO previously determined by mapping of the positions of the receptive fields (Boussaoud et al. 1991; Gattass et al. 1988). With the use of these borders the distribution of the ratio of S_{max}/MAX, which was introduced to quantify the distinctiveness of the selectivity to complex features, and that of the square root of the area of the receptive fields are compared for the four areas in Fig. 10. The distribution of the ratio of S_{max}/Max shifted toward 0 step by step from V2 to anterior IT. Although the difference between V4 and posterior IT was only numerical ($P > 0.05$), all the other pairs showed significant differences ($P < 0.001$). The size of the receptive field increased step by step from V2 to anterior IT, i.e., V2 < V4 < posterior IT < anterior IT. Because only the sample in V4 contained cells whose eccentricity exceeded 8° (open areas in Fig. 10,

right) we eliminated them from the comparison. All the pairs showed significant differences ($P < 0.001$).

Some of the complex critical features identified in the present study are listed in Fig. 11. These are selected because we have the greatest confidence that the feature could not be further simplified. The features in V4 and posterior IT varied as much as those in anterior IT. The features in V4 and those in posterior IT both included complex shapes, combinations of a shape and texture, and combinations of a shape and color. Some of the features in V4 and posterior IT appeared to be as complex as the critical features in anterior IT, although it is important to note that there exists no objective method for quantifying the complexity of pictorial features.

Most V2 cells were maximally activated by some simple features, that is, we failed to find a complex stimulus that elicited a response stronger than that to simple stimuli. However, three cells showed significantly stronger responses to particular complex features than to any simple stimuli. Because such cells have not been reported in V2, we will introduce two of them. We do not have confidence in the data of the third cell because the unit was lost before it had been tested with a sufficient variety of simple stimuli. The first cell, which is illustrated in Fig. 12, responded selectively to stimuli that contained patterns such as concentric rings (*a*). We found that a combination of three rings was necessary and sufficient to maximally activate the cell (*b-d*). Neither stripes of any orientation nor a combination of square-shaped rings was effective (*e-j*). The second cell, the data for which are shown in Fig. 13, responded maximally to a tapered bar (*b*). Although bars of various orientations and widths were carefully tested, the strongest response to bar was still 65% (*i*) of that to the tapered bar.

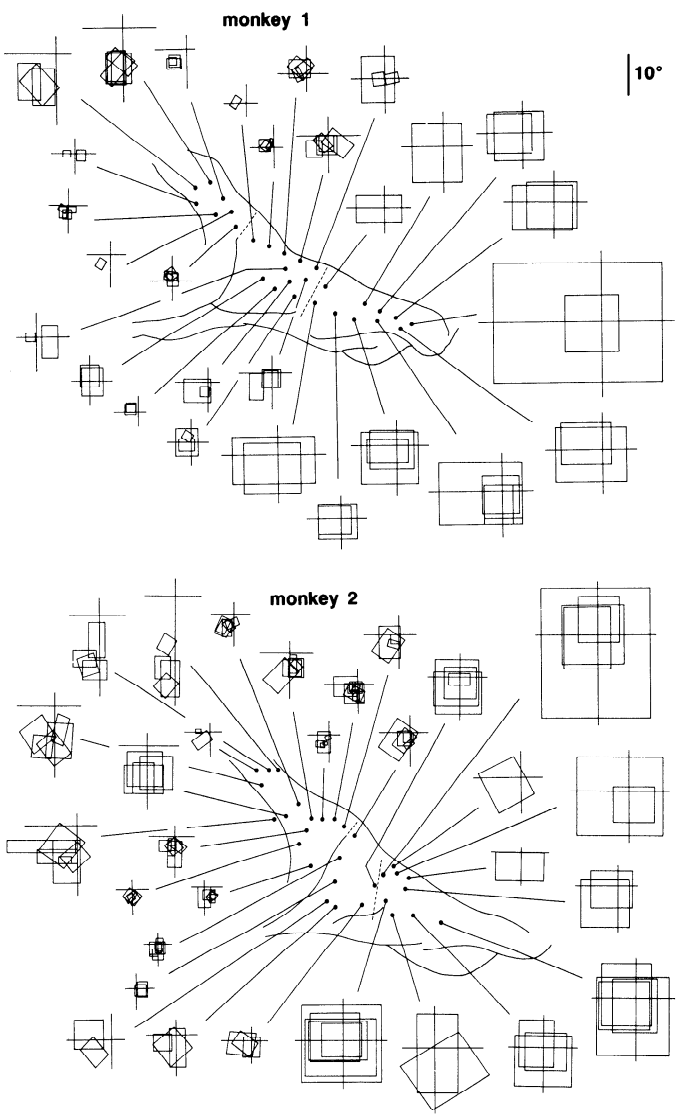

FIG. 9. Receptive fields. The receptive fields of cells recorded from a single penetration are superimposed. Broken lines: border between anterior IT and posterior IT and that between posterior IT and V4. The receptive fields in the penetration slightly posterior to the border between anterior IT and posterior IT in *monkey 2* were relatively large, but because of the selectivity of the cells (Fig. 8) and the position relative to the posterior middle temporal sulcus we positioned the penetration in posterior IT. The conclusions of this paper do not change at all if this penetration is moved to anterior IT. The receptive fields of V2 cells are not shown.

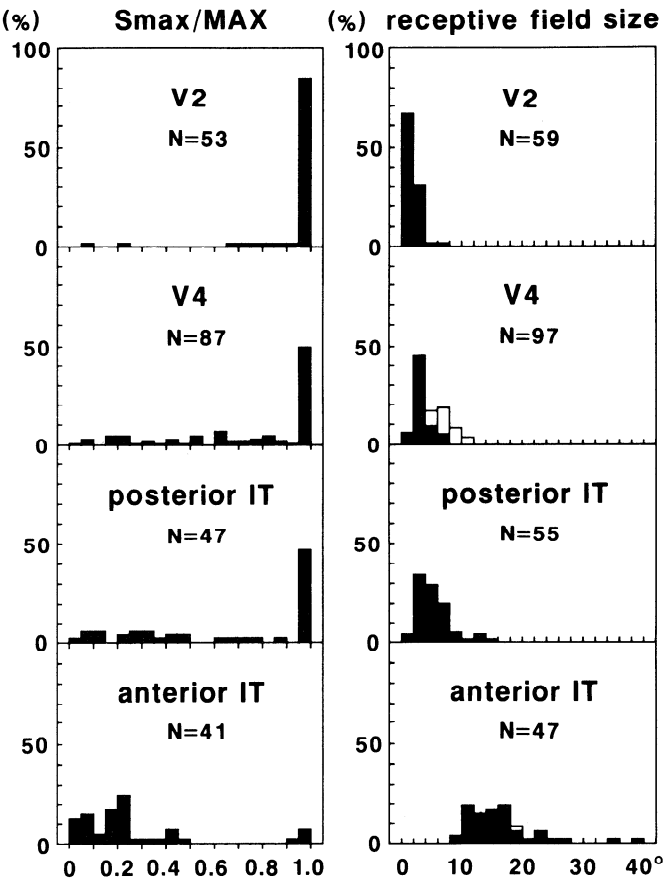

FIG. 10. Distribution histograms of the ratio of S_{max}/MAX and the size of the receptive field in the 4 regions. The size of the receptive field is given by the square root of the area of the receptive field. See METHODS for the method of determining the border of the receptive field and the method of calculation for the area. Filled areas in *right histograms*: cells having receptive fields with eccentricity <8°. Open areas: those having receptive fields with larger eccentricity. The means of the size of the receptive fields were 1.7 ± 1.0°, mean ± SD, in V2; 4.8 ± 2.5° for cells in V4; 5.4 ± 2.8° in posterior IT; and 16.5 ± 6.1° in anterior IT.

DISCUSSION

The results in this paper can be summarized in the following two points. *1*) A sizable proportion of cells in V4 and

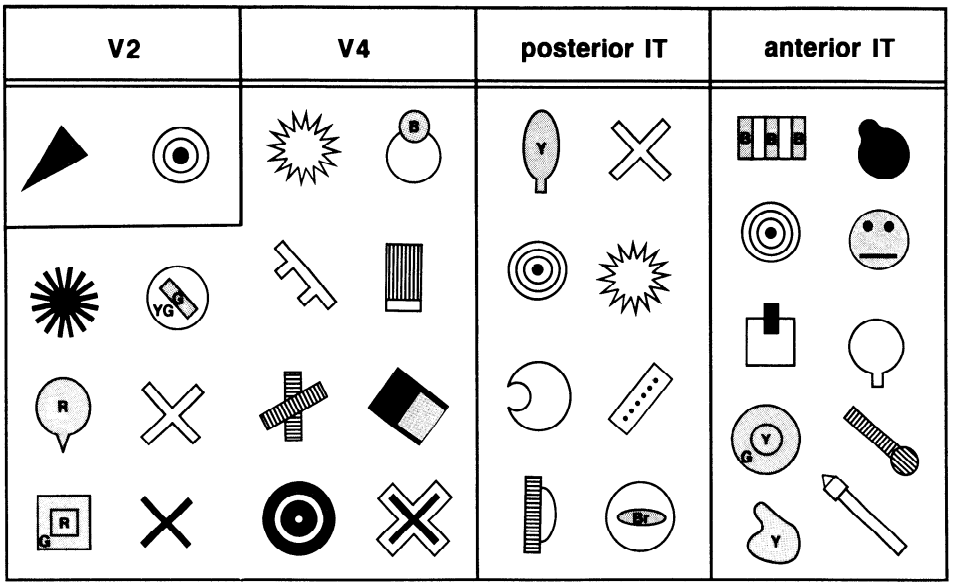

FIG. 11. Examples of the complex critical features in the 4 regions. YG, yellow green; Br, brown.

posterior IT, but not in V2, responded maximally to particular complex object features, as did cells in anterior IT. *2*) The selectivity of the cells in anterior IT was generally distinctive, but cells with selectivity of varying distinctiveness intermingled in V4 and posterior IT. Although the critical features for cell responses have previously been examined in anterior IT (Desimone et al. 1984; Gross et al. 1972; Tanaka et al. 1991), posterior IT, and V4 (Tanaka et al. 1991), the distinctiveness of the selectivity, which turned out to be a key issue in posterior IT and V4, was quantified for the first time in the present study. A prominent increase in the size of the receptive fields from posterior IT to anterior IT was also observed, but this is rather confirmation of the previous results (Desimone and Gross 1979; Tanaka et al. 1991).

One of the authors previously reported that only 2% of responsive cells in V4 and 12% of those in posterior IT selectively responded to complex object features (Tanaka et al. 1991). In the present study we found that greater proportion of cells in these areas, i.e., 38% in V4 and 49% in posterior IT ($S_{max}/MAX < 0.75$, Fig. 10), required complex features for their maximal activation. We found that many cells in these areas showed moderately strong responses to some simple stimulus as well as the maximum response to the complex critical feature. Such cells might be classified as "primary cells" in the previous study, in which the classification was performed mostly qualitatively by hearing discharges, because they responded to some simple stimuli. Also, the introduction of a special computer graphic system in the present experiments facilitated the exploration of effective stimuli and the quantitative comparison of the effectiveness of different stimuli.

Earlier, Tanaka et al. (1986) reported the presence of cells in the prelunate gyrus that specifically responded to "stimuli with an irregular internal structure or texture." Recently, Gallant et al. (1993) reported that a sizable proportion of V4 cells responded to concentric or hyperbolic patterns more strongly than to straight gratings of any orientation. Some of our critical features in V4 and posterior IT were similar to the concentric or hyperbolic patterns, but our critical features were more divergent.

At what point is the selectivity to complex object features attained? The distribution of S_{max}/MAX, the ratio of the maximum response to simple stimuli to the total maximum response of the individual cells, showed the most prominent changes from V2 to V4 and from posterior IT to anterior IT (Fig. 10). This might appear as evidence that integration of features advances in V4 and anterior IT. However, there is no reason to assume that the sample of cells represented outputs of the areas. Rather, the cells, which were randomly sampled at various depth, should have included cells at various stages of the local networks.

The ratio S_{max}/MAX showed the greatest intra-areal variety in posterior IT and V4 (Figs. 8 and 10). If we assume that the selectivity is determined in local networks but not in corticocortical connections, a random sample of cells from one such local network should include cells with varying complexity of selectivity. A cell located close to the input end should respond maximally to some primary feature that corresponds to a component of the final feature, a cell located close to the output end should respond rather selectively to the final integrated feature, and a cell located at the middle should show an intermediate property. The samplings from V4 and posterior IT but not those from anterior IT fulfilled this condition. The initial assumption is plausible on the basis of the complicated intrinsic connections in the local regions of the cerebral cortex (Lorente de Nó 1938; Lund 1988). The computational power of these complicated local circuitries should be greater than that of the one-step corticocortical connections. Thus we suggest that signals of primary features are integrated to form complex features in local networks of V4 and posterior IT.

The receptive fields of cells in V4 and posterior IT were smaller than those of cells in anterior IT. These small receptive fields are advantageous for integration of components because activities of cells with small receptive fields provide information regarding the position of the components, which may be necessary for the integration. If the integration occurred with large receptive fields such as those in anterior IT, there should be some sophisticated mechanism to register positional relationship between the components. On the other hand the large receptive fields in anterior IT

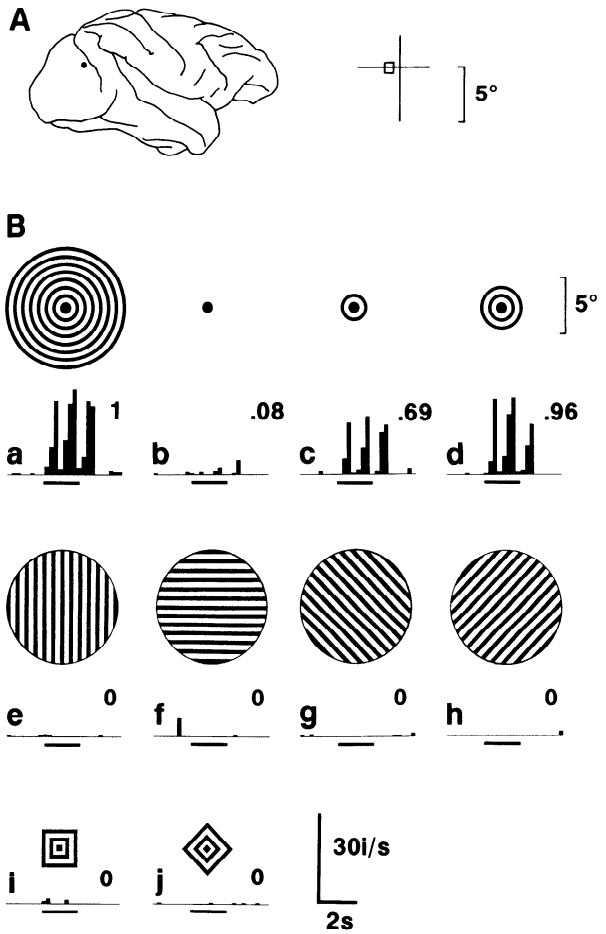

FIG. 12. V2 cell that responded exclusively to stimuli that contained concentric rings. A combination of 3 rings was necessary for maximal activation (*b-d*). The response in *c* was significantly weaker than those in *a* and *d* ($P < 0.01$). Straight stripes at any orientation (*e-h*) or combinations of square-shaped rings (*i* and *j*) were not effective at all. The 3 peaks in the responses reflected the circular movements of the pictures. There were $2\frac{1}{3}$ periods of movements during the 2-s stimulus presentation.

are useful for achieving the ultimate objective of the system, that is, position-independent discrimination of objects. A possibility is that the large receptive fields of anterior IT cells are realized by a converging connection from multiple sites in posterior IT with the same selectivity to a single site in anterior IT. But this means that each site in anterior IT has a specific input source in each retinotopical component of posterior IT, which may conflict with the fact that the peripheral visual field is less represented than the central visual field in posterior IT (Boussaoud et al. 1991). It is also possible that the selectivity to complex features develops in different ways in the central and peripheral visual fields and that for stimuli in the peripheral field it is established only in anterior IT. We are open to this possibility because we did not examine posterior IT cells having receptive fields in the peripheral visual field.

The concept that the basis for the selectivity is mostly situated in V4 and posterior IT is also consistent with anatomic results from this laboratory (Saleem et al. 1993). They made punctual injections of *Phaseolus vulgaris* leucoagglutinin (PHAL) into the lateral part of posterior IT, where the central visual field is represented, and found only a few dense terminal foci in anterior IT. If the output of posterior IT conveyed signals regarding primitive features, they should terminate in a greater number of sites in anterior IT. Primitive features should be used as components of many complex features. The labeled terminal foci were not limited to the middle layers but extended to all the cortical layers, forming a columnar shape. A previous physiological study from this laboratory showed that there is columnar organization in anterior IT, that is, cells with similar selectivity form a cluster elongated vertical to the cortical surface (Fujita et al. 1992). The columnar terminals of afferents from posterior IT may directly convey signals selective to complex features to cells in all the layers in the columnar region of anterior IT.

If the selectivity exists before signals reach anterior IT, then what is the function of anterior IT? We suggest that

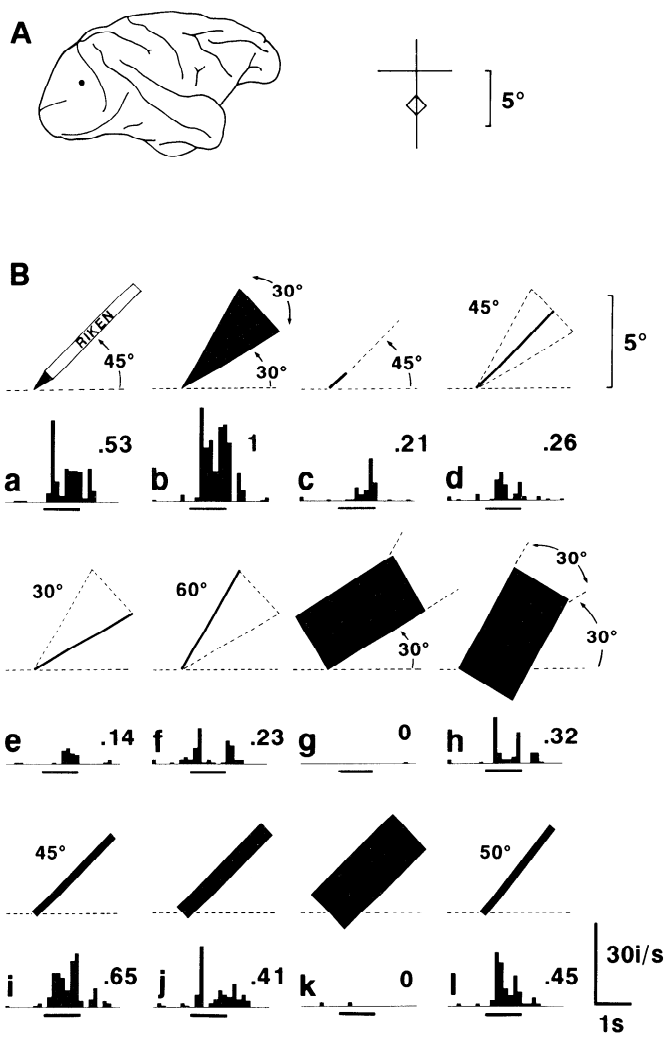

FIG. 13. V2 cell that responded maximally to a tapered bar. The cell responded strongly to a pencil (*a*) with distinctive selectivity for the orientation. A triangle with the same orientation and taper as those of the tip of the pencil evoked a stronger response (*b*). Either edge represented by a narrow line (*e* and *f*) or the edge of a wide rectangle (*g* and *h*) was much less effective. The width was found to be somewhat critical when tested with bars oriented between the edges (*d* and *i-k*) and 65% of the maximum response was evoked by a bar 0.5° in width (*i*). The orientation of the bar of this width was carefully changed but we could not obtain a stronger response (for example, *l*). All the other responses were significantly weaker than the response in *b* ($P < 0.01$).

selectivity is not the final goal of computation. The selectivity and the columnar organization may be tools used to compute something more. Intercolumnar and intracolumnar neuronal communication in anterior IT should be studied.

Although we found integration of features in only a few cells in V2, we would not dispute the role of V2 in the formation of selectivity to complex features. It is possible that selectivity to integrated features may be very moderate in V2. For example, a cell may respond to an integrated feature only 20% more strongly than to a simple stimulus. Or selectivity may appear only in a disinhibitory manner. Such moderate selectivity was not detected in the present study. It is also possible that circuitries in V2 are mainly used for the completion of primary features from the images [corresponding to the step from the raw primary sketch to the full primary sketch in the framework of Marr (1982)] rather than detection of complex features. It has been found that a sizable proportion of V2 cells responded to illusory contours in addition to ordinary contours (Peterhans and von der Heydt 1989; von der Heydt and Peterhans 1989; von der Heydt et al. 1984).

The authors thank I. Fujita and M. Ito for critical reading of the manuscript.

This work was supported by the Frontier Research Program, RIKEN, and grants from MESC of the Japanese government. E. Kobatake was supported by a postdoctoral fellowship from SRBSP, Japan.

Address for reprint requests: K. Tanaka, Laboratory for Neural Information Processing, Frontier Research Program, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-01, Japan.

Received 12 July 1993; accepted in final form 2 November 1993.

REFERENCES

ALBRECHT, D. G., DE VALOIS, R. L., AND THORELL, L. G. Visual cortical neurons: are bars of gratings the optimal stimuli? *Science Wash. DC* 208: 88–90, 1980.

BARLOW, H. B., BLAKEMORE, C., AND PETTIGREW, J. D. The neural mechanism of binocular depth discrimination. *J. Physiol. Lond.* 193: 327–342, 1967.

BOUSSAOUD, D., DESIMONE, R., AND UNGERLEIDER, L. G. Visual topography of area TEO in the macaque. *J. Comp. Neurol.* 306: 554–575, 1991.

DEAN, P. Effects of inferotemporal lesions on the behavior of monkeys. *Psychol. Bull.* 83: 41–71, 1976.

DESIMONE, R., ALBRIGHT, T. D., GROSS, C. G., AND BRUCE, C. Stimulus selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* 4: 2051-2062, 1984.

DESIMONE, R. AND GROSS, C. G. Visual areas in the temporal cortex of the macaque. *Brain Res.* 178: 363–380, 1979.

FELLEMAN, D. I. AND VAN ESSEN, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cerebr. Cortex* 1: 1–47, 1991.

FUJITA, I., TANAKA, K., ITO, M., AND CHENG, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature Lond.* 360: 343–346, 1992.

GALLANT, J. L., BRAUN, J., AND VAN ESSEN, D. C. Selectivity for polar, hyperbolic, and cartesian gratings in Macaque visual cortex. *Science Wash. DC* 259: 100–103, 1993.

GATTASS, R., SOUSA, A. P. B., AND GROSS, C. G. Visuotopic organization and extent of V3 and V4 of the macaque. *J. Neurosci.* 8: 1831–1845, 1988.

GROSS, C. G. Visual functions of inferotemporal cortex. In: *Handbook of Sensory Physiology,* edited by R. Jung. Berlin: Springer-Verlag, 1972, p. 451–482.

GROSS, C. G., BENDER, D. B., AND ROCHA-MIRANDA, C. E. Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science Wash. DC* 166: 1303–1306, 1969.

GROSS, C. G., ROCHA-MIRANDA, C. E., AND BENDER, D. B. Visual properties of neurons in inferotemporal cortex of the macaque. *J. Neurophysiol.* 35: 96–111, 1972.

IWAI, E. AND MISHKIN, M. Extrastriate visual focus in monkeys: two visual foci in the temporal lobe of monkeys. In: *Neurophysiological Basis of Learning and Behavior,* edited by N. Yoshii and N. A. Buchwald. Osaka, Japan: Osaka Univ. Press, 1968, p. 23–30.

KOBATAKE, E. AND TANAKA, K. Selective neuronal responses to complex visual object features are already present in posterior part of the macaque inferotemporal cortex. *Soc. Neurosci. Abstr.* 17: 443, 1991.

KOBATAKE, E. AND TANAKA, K. Selectivity for features beyond orientation, color, size, and simple texture in the prestriate areas V2 and V4. *Soc. Neurosci. Abstr.* 18: 146, 1992.

LORENTE DE NÓ, R. The cerebral cortex: architecture, intracortical connections and motor projections. In: *Physiology of the Nervous System,* edited by J. F. Fulton. New York: Oxford Univ. Press, 1938, p. 291–339.

LUND, J. S. Anatomical organization of macaque monkey striate visual cortex. *Annu. Rev. Neurosci.* 11: 253–288, 1988.

MAGUIRE, W. M. AND BAIZER, J. S. Visuotopic organization of the prelunate gyrus in rhesus monkey. *J. Neurosci.* 4: 1690–1704, 1984.

MARR, D. *Vision.* New York: Freeman, 1982.

MISHKIN, M. A memory system in monkey. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298: 85–95, 1982.

PETERHANS, E. AND VON DER HEYDT, R. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J. Neurosci.* 9: 1749–1763, 1989.

SALEEM, K. S., TANAKA, K., AND ROCKLAND, K. S. Specific and columnar projection from area TEO to TE in the macaque inferotemporal cortex. *Cerebr. Cortex.* 3: 454–464, 1993.

SCHWARTZ, E. I., DESIMONE, R., ALBRIGHT, T. D., AND GROSS, C. G. Shape recognition and inferior temporal neurons. *Proc. Natl. Acad. Sci. USA* 80: 5776–5778, 1983.

TANAKA, K., SAITO, H., FUKADA, Y., AND MORIYA, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* 66: 170–189, 1991.

TANAKA, M., WEBER, H., AND CREUTZFELDT, O. D. Visual properties and spatial distribution of neurones in the visual association area on the prelunate gyrus of the awake monkey. *Exp. Brain. Res.* 65: 11–37, 1986.

VON DER HEYDT, R., PETERHANS, E., AND BAUMGARTNER, G. Illusory contours and cortical neuron responses. *Science Wash. DC* 224: 1260–1262, 1984.

VON DER HEYDT, R. AND PETERHANS, E. Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J. Neurosci.* 9: 1731–1748, 1989.

VON DER HEYDT, R., PETERHANS, E., AND DURSTELER, M. R. Periodic-pattern-selective cells in monkey visual cortex. *J. Neurosci.* 12: 1416–1434, 1992.