Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns

Kazushige Tsunoda[1,2], Yukako Yamane[1,3], Makoto Nishizaki[1] and Manabu Tanifuji[1]

[1] *Laboratory for Integrative Neural Systems, Brain Science Institute, The Institute of Physical and Chemical Research (RIKEN), Hirosawa 2-1, Wako-shi, Saitama 351-0198, Japan*

[2] *Present address: Department of Ophthalmology, Ashikaga Red Cross Hospital, Honjo 3-2100, Ashikaga-shi, Tochigi 326-0808, Japan*

[3] *Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan*

Correspondence should be addressed to M.T. (tanifuji@postman.riken.go.jp)

Intrinsic signal imaging from inferotemporal (IT) cortex, a visual area essential for object perception and recognition, revealed that visually presented objects activated patches in a distributed manner. When visual features of these objects were partially removed, the simplified stimuli activated only a subset of the patches elicited by the originals. This result, in conjunction with extracellular recording, suggests that an object is represented by a combination of cortical columns, each of which represents a visual feature (feature column). Simplification of an object occasionally caused the appearance of columns that were not active when viewing the more complex form. Thus, not all the columns related to a particular feature were necessarily activated by the original objects. Taken together, these results suggest that objects may be represented not only by simply combining feature columns but also by using a variety of combinations of active and inactive columns for individual features.

Information about visually presented objects is transmitted from the primary visual cortex (V1) to the inferotemporal cortex (IT) through multiple pre-striate areas in macaque monkeys. The anterior part of the IT cortex, defined architechtonically as area TE, is the final exclusively visual area in this pathway (ventral visual pathway), and has strong connections with medial temporal structures essential for the formation of recognition memory. To understand the neural mechanisms of perception and recognition of objects by their visual images, visual responses of individual neurons in this area have been extensively investigated by extracellular recordings. Although there is a class of neurons that responds specifically to the sight of complex objects such as faces or hands[1–3], this approach has shown that most neurons in this area respond optimally to the geometrically less complex features than to more complex real objects[3–5]. These findings suggest that activation of single TE neurons is not necessarily enough to specify given objects, and that activation of multiple neurons is involved in the neural representation of an object. How objects are represented by the activation of multiple neurons, however, is still uncertain, despite the basic importance of this issue.

Intrinsic signal imaging is a promising technique for observing the activation of multiple neurons at the same time, and has been successfully used in various areas including area TE[6–14]. It has the advantage over electrophysiological recordings that the local average of neuronal activities can be measured simultaneously over a wide cortical region. Here, to further investigate the representations of complex objects in dorsal TE, we applied this technique to examine the spatial patterns of activation evoked by various complex objects and simplified forms of them. In addition, stimulus simplification reveals systematic changes in the spatial patterns of neural activation, which suggests an extended model of the feature-based representation; namely, objects are represented not only by simply combining feature columns but also by using a variety of combinations of active and inactive columns for individual features.

Results

Intrinsic signal imaging revealed multiple spots elicited by visual stimulation in dorsal area TE. The mean (± s.d.) size of each 'active spot' (**Fig. 1g**) was 0.50 ± 0.13 mm along the longer axis and 0.35 ± 0.09 mm along the shorter axis (n = 94). (See Methods for identification and demarcation of active spots.) These dimensions corresponded with the size previously reported for a column of cells with similar responsiveness in this area[15,16]. The location and shape of individual spots' contours were essentially the same in different threshold levels (**Fig. 1h**). Furthermore, changes in the statistical significance levels did not affect the identification of active spots (**Fig. 1b–d**).

To relate the optical response specificity to the neural activities, we recorded extracellular activity from 34 neurons in 17 sites located inside and outside the active spots (**Fig. 2a**). The representative peristimulus-time histograms (PSTHs; **Fig. 2b**) showed that the neuron in site 2, where both stimulus 1 and 2 elicited significant intrinsic signals, was activated by these 2 stimuli, but not by stimulus 3. Similarly, only stimulus 1 significantly activated the cell in site 5 and none of the stimuli activated cells in site 6. In summary, among 34 neurons, 28 cells (82.4%) showed the same responsiveness to stimulus 1 as indicated by the optical responses: 8 active cells inside the spot and

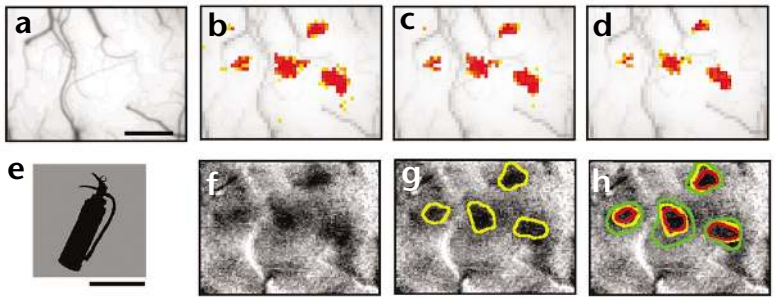


Fig. 1. Intrinsic signal imaging detects local modulation of light absorption changes in area TE. (**a**) Surface view of the exposed portion of dorsal area TE. (**b–d**) Active regions, elicited by the visual stimulus shown in (**e**), in which the reflection change elicited by the stimulus was significantly greater than that elicited by the control. The highest significance level is denoted by red, the lowest by yellow (**b**, $p < 0.05$; **c**, $p < 0.025$; **d**, $p < 0.01$). (**f**) A filtered differential image showing local increase in absorption. (**g**) Extracted active spots outlined by connecting pixels with half the peak absorption value. (**h**) Active spots outlined by different thresholds; 1/3 (red), 1/2 (yellow) and 2/3 (green) of the peak absorption value. Scale bars, 1.0 mm (**a–d**, **f–h**); 10° (**e**).

20 inactive cells outside (**Fig. 2c**, top). Similar results were obtained for 32 of 34 cells (94.1%) for stimulus 2 (**Fig. 2c**, middle), and for 33 of 34 cells (97.1%) for stimulus 3 (**Fig. 2c**, bottom). Throughout the investigation, we examined the correspondence between extracellular and optical responses in three cases (**Table 1**). Altogether, we found that 81.7% of neurons inside the spots were significantly activated by the visual stimuli specific for the spots, and 84.3% of neurons outside the spots were not responsive to the stimuli. A large variation of neuronal responses within the active spots (for example, **Fig. 2c**) did not affect the agreement between the intrinsic signals and extracellular responses. First, the optical response intensity and the firing rates of individual neurons showed statistically significant positive correlation ($p < 0.001$) (**Fig. 2d**). Second, the average of extracellular responses inside the spots (mean ± s.d., 24.9 ± 17.2 and 16.8 ± 11.6 spikes/s for stimulus 1 and 2, respectively) was significantly larger than the average of responses outside the spots (4.30 ± 4.95 and 3.32 ± 4.72 spikes/s for stimulus 1 and 2, respectively; **Fig. 2c**). These results confirm that the intrinsic signals coincide well with the activity of neurons examined by conventional extracellular recordings.

Active spots elicited by complex objects showed different distribution patterns for different objects (**Figs. 2a** and **3a**). At the same time, however, some spots activated by one object were also activated by other objects. These observations made us consider that a spot represented a visual feature, and thus the overlap of spots was observed when the feature was common among the objects. We examined this possibility by comparing active spot distribution patterns produced by complex stimuli and by systematically simplified stimuli (**Figs. 3b–d** and **6**). In Fig. 3b, a 'black cat' (**Fig. 3b-1**) was simplified to the 'head' (**Fig. 3b-2**), and then to the 'silhouette of the head' (**Fig. 3b-3**). The original image elicited 14 spots, but presenting the head elicited only 8 of the original 14 spots. The silhouette only activated three (arrows) of the eight spots elicited by the head. For 12 pairs of activation patterns elicited by the original and the simplified stimuli, 5 pairs (42%) showed similar results: simplified stimuli lacking some of the visual features of the original image activated only a subset of the spots elicited by the stimuli before simplification. These results were consistent with the idea that individual columns represent visual features rather than object images.

In addition to the spot disappearance by simplification, however, there were cases in which new spots emerged by an apparent simplification of an object (**Fig. 3c** and **d**, **Fig. 6**). In one case (**Fig. 3c**), the simplified stimulus (**Fig. 3c-2**) activated four new spots (arrows) in addition to two of three spots activated by the original object (**Fig. 3c-1**). Similarly, in another case (**Fig. 3d**), spots A and B disappeared, but spot C appeared when stimulus 1 was simplified to stimulus 3. Among the 12 pairs of activation patterns elicited by original and simplified stimuli, the emergence of spots by simplification was observed in 7 pairs (58%). (Two cases showed only emergence, and five cases showed both emergence and disappearance.)

One possible interpretation of spot emergence following simplification is that, at least for some features, only part of the spots representing a particular feature were activated when the feature was embedded in a complex object. For example, in the trial shown in Fig. 3d, both spots C and D seem to be related to the 'cylinder of the fire extinguisher' (**Fig. 3d-3**), but only spot D was activated by the original 'fire extinguisher' (**Fig. 3d-1**). We examined this possibility by single-cell recordings from the individual spots in two cases (**Figs. 3d** and **6**).

Twenty-five cells from 16 penetration sites were recorded in the case shown in Fig. 3d. We first examined neuronal responses to the visual stimuli used for optical imaging (quantitatively, or for some cases, manually), and confirmed that response selectivity was consistent with the previous optical imaging (**Table 1**). We then analyzed in more detail the responsiveness of these cells from each spot (**Figs. 4** and **5**). The difference in optical response patterns to stimulus 1 and 3 in Fig. 3d suggested that spots A and B represented visual features related to 'handle and hose' of the fire extinguisher. In fact, two cells in spot A and three cells in spot B were significantly activated by the handle and hose in isolation (**Fig. 4a-2** and **b-2**) as well as by the silhouette of the original fire extinguisher (**Fig. 4a-1** and **b-1**). The cells in spot A were activated by the handle (**Fig. 4a-3**) having protrusions, but not by the hose (**Fig. 4a-4**). Furthermore, other stimuli with sharp protrusions, such as a hand (**Fig. 4a-5**) and cat's head (**Fig. 4a-6**), also activated the cells. These cells seemed to require sharp protrusions for activation (**Fig. 5a**). In contrast, three cells in spot B were activated by the hose (**Fig. 4b-4**), but neither by the handle (**Fig. 4b-3**) nor a line segment (**Fig. 4b-5**). We determined the feature critical for two of the cells as an asymmetric arc (**Fig. 5b**). The remaining one cell was significantly activated by the asymmetric arc, but the strongest responses were elicited by a banana, which also has curved shape (**Fig. 5b**). When the hand and cat's head were manually presented, there was no activation in these three cells.

The cylinder (**Fig. 3d-3**) produced significant optical responses in both spots C and D, but other stimuli having the cylinder as a part (**Fig. 3d-1**, **d-2** and **d-4**) only activated the spot D. The neural responses of cells in these spots were consistent with these imaging results: a cell in spot C was activated by the cylinder but not by the original fire extinguisher (**Fig. 4c-1** and **c-2**), and a cell in spot D was significantly activated by both stimuli (**Fig. 4d-1** and **d-2**). The feature critical for the cell in spot D was a rectangular shape (**Fig. 4d-3**), but the cell also responded significantly to an ellipse (**Fig. 4d-4**). Because there was no response to a circle (**Fig. 4d-5**), an elongated structure seemed to be necessary for

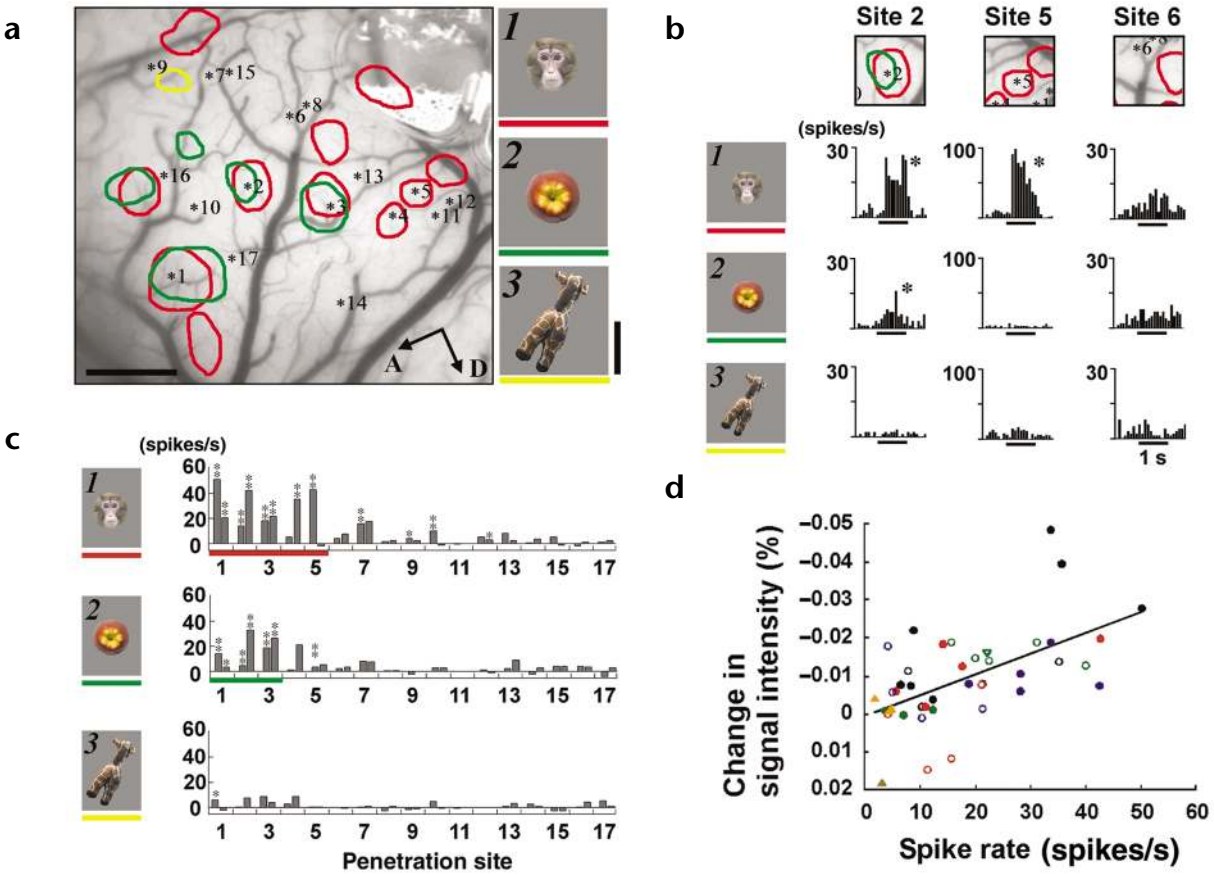


Fig. 2. Relationship between intrinsic signals and spike activity in area TE. (**a**) Active spots elicited by three different stimuli (1, 2 and 3, right) and numbered electrode penetration sites. The color of individual contours indicates the active spots elicited by the stimulus underlined with the same color (same in all the figures showing spot distribution). A, anterior; D, dorsal. Horizontal scale bar (black), 1.0 mm (**a**); vertical scale bar, 10° (applies to insets). (**b**) Representative PSTHs showing the extracellular activity elicited by three different stimuli, recorded at the sites indicated at the top of each column. Each row gives the PSTHs obtained by the stimuli shown at the left. Horizontal scale bars (black), one-second periods of visual stimulation. (**c**) Mean firing rates evoked by the 3 different visual stimuli, for all 34 cells from 17 different sites. Two different cells separated by at least 200 μm were recorded at each penetration site. Colored bars in the top (red) and middle graphs (green), penetration sites inside the active spots of the stimulus. For stimulus 3, all the cells were recorded outside the active spot. (**d**) Correlation between the intensity of local changes of optical signal and evoked spike rates measured at the same cortical locations. Thirty-four cells were tested with 10 stimuli, and cellular activity with significant visual responses ($p < 0.05$) was plotted against the optical signal intensity obtained at the corresponding penetration sites for the individual stimuli. Activity of neurons from the same sites was averaged. Different symbols indicate the responses at the different sites. The regression line is given by $y = -5.05e^{-6} + 5.46e^{-6}x$ with a correlation coefficient of 0.57 ($n = 41$). Significance of Kolmogorov–Smirnov test for individual cell response, $^{*}p < 0.05$, $^{**}p < 0.01$.

activation. Among eight cells recorded from spot D, seven required an elongated structure for activation. The critical features were the ellipse and rectangular shape for one and three cells, respectively. Three of the remaining cells were activated maximally by a variety of stimuli also having elongated structures. Only one cell responded maximally to a circle (**Fig. 5d**).

For 9 of 12 cells recorded from spot C, the simplest visual feature was a rectangular shape (**Fig. 4c-3**) or rectangles with slightly rounded corners (**Fig. 5c**). In contrast to the cells in spot D, none of the 12 cells was activated by an ellipse (**Fig. 4c-4**). In addition, all of these cells were inhibited by the circle (six were significant; **Fig. 4c-5** and **Fig. 5c**). These results suggest that the response properties of the cells in spot C (**Fig. 3d**) were determined by the balance between excitatory and inhibitory inputs; that is, the excitatory inputs were given by a feature related to a rectangular shape and the inhibitory inputs were given by a feature related to a circular shape. This explanation would account for the lack of activation by the fire extinguisher, for which a circular hose was attached to a rectangular cylinder.

These results suggest that some of the columns representing a particular feature were inactive when other features were presented together with the feature. This could explain the optical imaging results in which active spots appeared following simplification of stimuli. This is supported by similar results in another case (**Figs. 6** and **7**). Here, the optical imaging showed that two spots, including spot A, were activated exclusively by the 'padlock without an arm' (stimulus 2), although the other three spots including spot B were activated both by the original 'padlock' (stimulus 1) and by the 'padlock without an arm' (**Fig. 6**). We recorded 14 cells from spot A. Most of them were significantly activated by the padlock without an arm (11 of 14 cells) and only a few cells were activated by the original padlock (2 of 14 cells; **Fig. 7a**). Twelve of them were further examined by the image of an isolated arm. Seven cells were inhibited by the isolated arm (three were significant, $p < 0.05$; **Fig. 7a**). The rest of the cells were not responsive to the stimulus. Because spot A was activated by the padlock without arm but inhibited by the arm itself, the original padlock showed no activation in

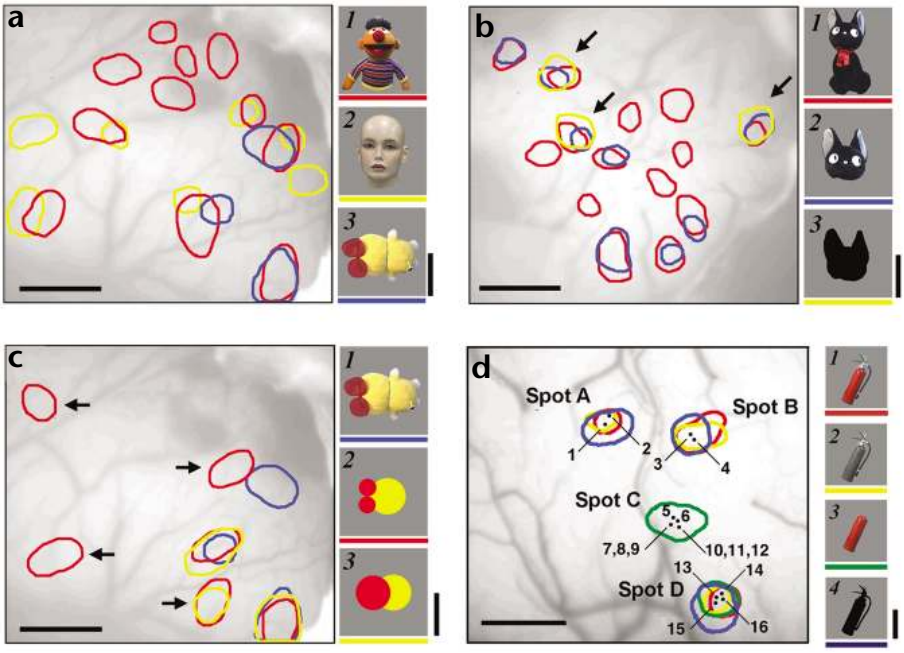


Fig. 3. Representation of complex object images and simplifications of them in area TE. (**a**) Spatial distributions of active spots elicited by three different objects. (**b**) A case in which simplified stimuli elicited only a subset of the spots evoked by the more complex stimuli. Arrows, common spots among the three stimuli. (**c**, **d**) Cases in which new spots appeared when the original stimulus was simplified (arrows in **c**). (**d**) Numbers (1 to 16) indicate electrode penetration sites. Horizontal scale bars (black), 1.0 mm; vertical scale bars, 10° (apply to all insets).

spot A. We recorded eight cells from spot B. In contrast to cells in spot A, all but one cell were significantly activated both by the original padlock and the padlock without an arm, but not by the isolated arm (**Fig. 7b**). (The one exception was exclusively activated by the padlock without an arm.)

Discussion

Here we demonstrate that combination of intrinsic signal imaging and extracellular recording can reveal the spatial layout of neural activity evoked by complex objects. This provides further opportunity to consider the neural mechanism of object perception and recognition. We show that objects are represented by combination of multiple columns in a sparsely distributed manner. The region activated by a single object image is only 3.3 ± 2.5% of the entire recording area (number of examined object images, 37). We also show that object simplification results in systematic changes in the distributed patterns of columns.

In general, two models have been proposed for object representation. One proposal is that an object is represented by a combination of modules, each specific to a visual feature or a part of the object (feature-based or part-based representation)[4,17–19]. An alternate proposal is that modules are specific to the object (object-based representation). In this paper, we showed that the simplified stimuli lacking some of the visual features activated only a subset of the spots elicited by the non-simplified stimuli in 42% of cases, and that optimal stimuli for neurons in these spots were visual features less complex than the original objects. These results are consistent with feature-based representation, in which an object is represented by the combined activation of columns, each responding to a specific visual feature of the object. In support of this interpretation, Fig. 6 shows that a red cup, which is different from the padlock as a visual image, activates spot B, as expected from response properties of cells in spot B (**Fig. 7**). The representation of objects by combinations of feature columns is further supported by previous reports showing that many neurons in this area respond better or equally well to features than to real objects that are geometrically more complex[3–5]. Finally, fMRI studies provide circumscribed evidence for the distributed representation of objects in the homologous areas in humans[20,21], although the limitation in spatial resolution of fMRI does not allow direct characterization of functional modules.

In contrast, however, emergence of spots by object simplification observed in 58% of cases cannot be explained by the conventional framework of feature-based representation, which implicitly assumes that all the columns related to a feature in a visual stimulus are activated. To explain our results by the feature-based model, we propose an extended scheme of feature-based representation, in which objects are represented not by the simple

Table 1. Number of neurons correlating with the intrinsic signals.

	Inside active spots	Outside active spots	Total
Case 1 (Fig. 2a)			
Stimulus 1	8/10 (80.0%)	20/24 (83.3%)	28/34 (82.4%)
Stimulus 2	6/6 (100.0%)	26/28 (92.9%)	32/34 (94.1%)
Stimulus 3	not recorded	33/34 (97.1%)	33/34 (97.1%)
Average	**(90.0%)**	**(91.1%)**	**(91.2%)**
Case 2 (Fig. 3d)			
Stimulus 1	10/13 (76.9%)	8/12 (66.7%)	18/25 (72.0%)
Stimulus 2	8/13 (61.5%)	5/6 (83.3%)	13/19 (68.4%)
Stimulus 3	13/20 (65.0%)	3/5 (60.0%)	16/25 (64.0%)
Stimulus 4	9/13 (69.2%)	4/6 (66.7%)	13/19 (68.4%)
Average	**(68.2%)**	**(69.2%)**	**(68.2%)**
Case 3 (Fig. 6)			
Stimulus 1	7/8 (87.5%)	12/14 (85.7%)	19/22 (86.4%)
Stimulus 2	19/22 (86.4%)	not recorded	19/22 (86.4%)
Stimulus 3	not recorded	20/20 (100.0%)	20/20 (100.0%)
Average	**(86.9%)**	**(92.6%)**	**(90.9%)**
Average of case 1, 2 and 3			
	(81.7%)	**(84.3%)**	**(83.4%)**

Ratio of number of neurons that showed significant response in the active spots (or no response outside the active spots) to the total number of recorded neurons.

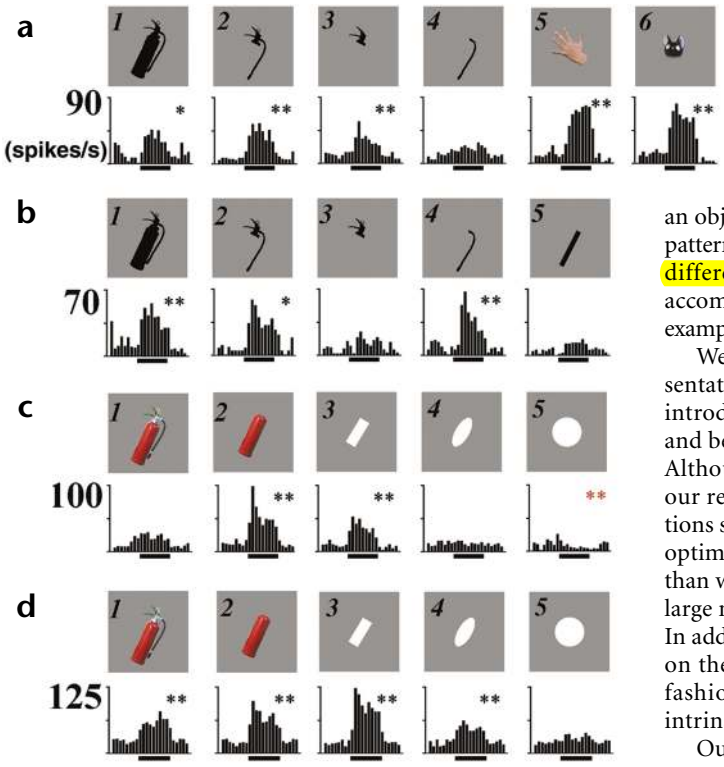


Fig. 4. Visual responsiveness of representative cells in spots A–D in Fig. 3d. (**a–d**) Responses in spots A (track 2; depth, 620 μm), B (track 3; depth, 540 μm), C (track 8; depth, 280 μm) and D (track 16; depth, 280 μm), respectively. Red asterisks indicate significant inhibition. *$p < 0.05$, **$p < 0.01$. Scale bars, one-second periods of stimulation.

sum of feature columns but by combinations of active and inactive columns for individual features. According to this scheme, we think that the fire extinguisher in Fig. 3d was represented not only by the active spots representing the handle (spot A), the hose (spot B) and the cylinder (spot D), but by the absence of spot C for representing the whole structure having elliptic shape. This extended model increases the number of available activation patterns, so that partial activation of columns can help in specifying an object. We speculate that the increased number of activation patterns allows for precisely differentiating different objects with different arrangements of features. Such combinations may accommodate changes in the form of an object that occur, for example, by rotation, partial occlusion, shading or changes in size.

We interpret our results based on the feature-based representation because this allows a consistent explanation without introducing extreme assumptions to the conventional model and because it is less speculative than the object-based model. Although we strongly favor the feature-based interpretation, our results by themselves do not exclude alternate interpretations such as the object-based representation. For example, the optimal stimuli for individual neurons could be more complex than what we found, especially because we cannot examine the large number of visual stimuli needed to refute this possibility. In addition, our analysis of the intrinsic signal imaging is based on the distribution of active spots defined in an all-or-none fashion, because of the limited signal-to-noise ratio of the intrinsic signals (**Fig. 1**).

Our extracellular recordings suggest that inhibitory mechanisms are involved in determining whether feature columns are active or inactive (**Figs. 4** and **7**). For example, the cells in spot A (**Fig. 6**) are inhibited by the arm of the padlock. In this case, spot A may act as an inactive column for the representation of the original padlock, but as an active column for the representation of the padlock without the arm. Similarly, cells in spot C (**Fig. 3d**) are inhibited by a circle, so that the spot may act as an inactive column for the representation of the fire extinguisher. One report suggests that intrinsic connections in area TE are involved in inhibitory interactions[22].

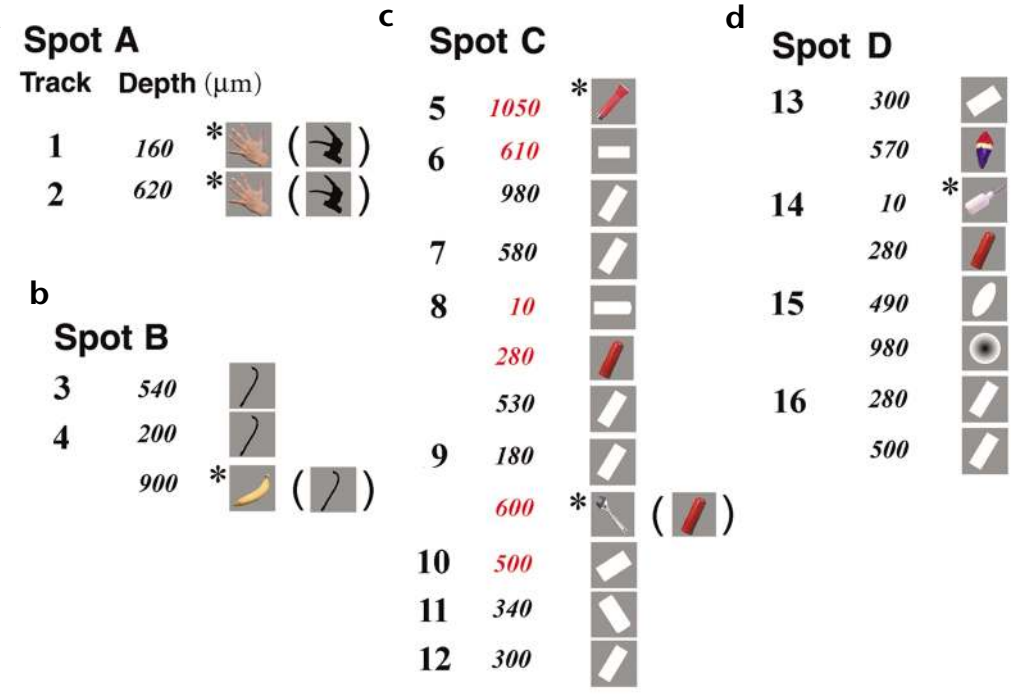


Fig. 5. The optimal visual stimuli of 25 cells in spots A to D in Fig. 3d. Stimuli without asterisks, visual features critical for the activation of the cells. For cells that showed maximal activation by the original objects among the original and simplified stimuli, original objects are indicated with asterisks, and visual features eliciting significant responses are indicated in parentheses. The track number, the same as the one indicated in Fig. 3d, and cortical depth are indicated to the left of the stimulus for each cell. All of these stimuli evoked significant responses ($p < 0.05$). Cortical depths indicated in red font represent cells significantly inhibited by a circle ($p < 0.05$).

Fig. 6. The distribution pattern of active spots elicited by two original object images (stimulus 1 and 3) and a simplified image of stimulus 1 (stimulus 2). The numbers (1 to 8) indicate electrode penetration sites. Horizontal scale bar (black), 1.0 mm; vertical scale bar, 10° (applies to insets).

In summary, our results from both optical and extracellular recordings provide evidence that a complex object is represented by combinations of columns in macaque area TE, and that the individual columns represent component visual features of an object. The combinations are not simply based on summing up the feature columns, but may instead rely on a combination of active and inactive columns to represent objects. However, the brain may use several mechanisms in parallel, such as feature-based and object-based representations. Some neurons in this area prefer the whole object configurations, such as faces and hands[1–3,12,13,23–25]. In addition, adjacent areas anterior to the recorded region are more related to visual recognition memory and may have a different form of object representation. When monkeys were extensively trained with real objects, landscapes and computer-generated complex features, neurons in more anterior part of inferotemporal cortices including perirhinal cortex responded to these trained visual stimuli[26,27].

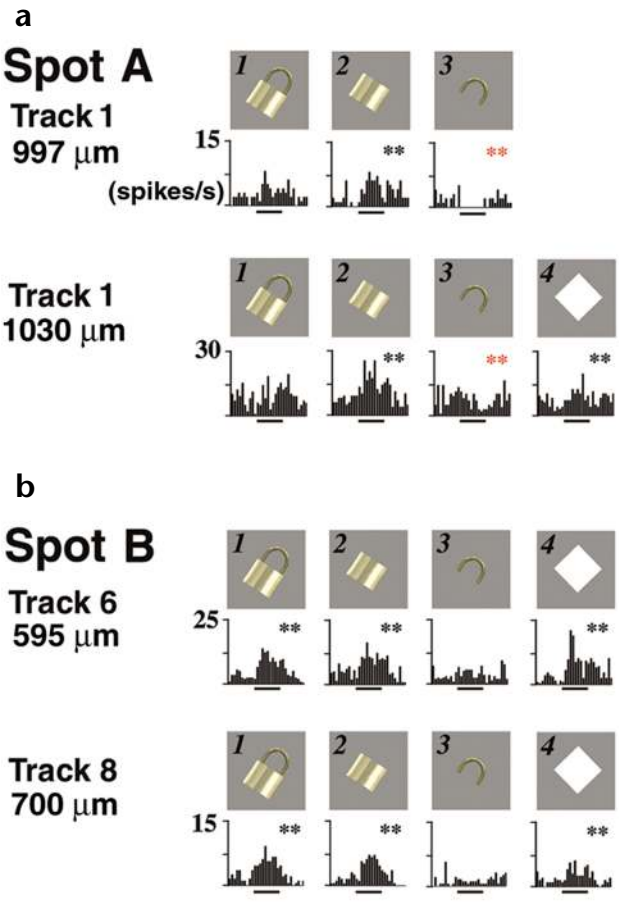
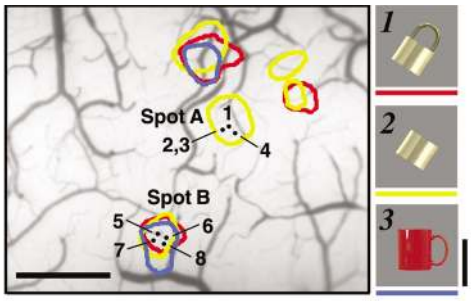


Fig. 7. Visual responsiveness of representative cells in spots A and B in Fig. 6. The PSTHs of two cells in spot A (**a**, track 1, depth, 997 μm and 1030 μm) and two cells in spot B (**b**; track 6, depth, 595 μm and track 8, depth, 700 μm) are shown. The fourth column gives responses to the simplest feature for these cells (white square). Red asterisks indicate significant inhibition. $^{**}p < 0.01$.



How feature-based and object-based representations are related remains the subject of future investigations.

Methods

Experimental procedure. The optical imaging and extracellular recording experiments were done on two hemispheres of two Japanese monkeys (*Macaca fuscata*) and three hemispheres of two Rhesus monkeys (*Macaca mulatta*) under anesthesia. The monkeys were paralyzed with pancuronium bromide or vecuronium bromide, and artificially ventilated with a mixture of N_2O, O_2 and isoflurane (70% N_2O, 30% O_2, isoflurane up to 1.0%). EEG, ECG, expired CO_2 concentration and rectal temperature were monitored throughout the experiments. The experimental protocol was approved by the Experimental Animal Committee of the RIKEN Institute. All experimental procedures were done in accordance with the guidelines of the RIKEN Institute and the National Institutes of Health.

Intrinsic signal imaging. The intrinsic signal (absorption changes at 605 nm of light) in the dorsal part of area TE was recorded through the glass coverslip window of a stainless steel chamber centered 15.0–17.5 mm anterior to the ear bar position[12,13]. Images were obtained using a video camera (CS8310, Tokyo Electronic Industry, Tokyo, Japan) and digitized with an IBM/PC-compatible equipped with a video frame grabber board (Pulsar, Matrox, Quebec, Canada)[14]. The imaged area was 8.8 × 6.6 mm and contained 320 × 240 pixels. Images of surface blood vessels were made under 540 nm light before the intrinsic signal imaging. The focusing depth was then adjusted to 350 to 500 μm below the cortical surface for the intrinsic signal imaging. Ten to eighteen different stimuli were randomly presented with a television monitor repeatedly 36 to 68 times for 2.0 or 4.0 s each, at 10- to 20-s intervals. The stimulus set included the original images of complex objects, their simplified images, and a control without stimulus. The simplification procedure included dividing the original image into separate parts and using one of them as a simplified stimulus, making an achromatic gray image of the original and filling the whole original image or particular parts of it with the color of the original, to mask visual features embedded in the original.

Data analysis. The stimuli evoked non-specific global darkening (increased absorption) of the cortex together with stimulus-specific local modulation of darkening. The active spots were extracted as follows. A differential image was obtained by dividing the averaged images obtained at 0.5 to 2.5 s after the stimulus onset by those obtained during a 1-s period before the stimulus onset. A Gaussian spatial filter was used to eliminate stimulus non-specific darkening and high-frequency noise of the differential image (cutoff frequencies, $\sigma = 20.4$/mm for high cut and $\sigma = 0.3$/mm for low cut). The pixel intensities of the images with stimulus were compared by *t*-test with those without stimulus (**Fig. 1b–d**). Intensity of local modulation was calculated by averaging the filtered images across all the trials for individual stimuli (**Fig. 1f**). The regions that showed significant darkening ($p < 0.025$) were defined as active spots, and outlined by connecting pixels with half the peak absorption value (**Fig. 1g**). Regions elongated along the cortical surface vessels were excluded from the analysis.

Electrophysiological recording. For extracellular recordings, the exposed cortex used for the optical imaging was covered with a transparent arti-

ficial dura made of silicone rubber (originally designed by A. Arieli[28]). Tungsten microelectrodes were inserted into the cortex through the artificial dura at sites determined relative to the surface blood vessel pattern. The depth of the recording sites was measured by a micromanipulator in distance from the cortical surface. For all the extracellular recordings in this study, spike responses were averaged over 10 or 20 stimulus presentations, and the difference in the firing rate during and before the stimulus presentation was examined by the Kolmogorov–Smirnov test ($p < 0.05$ or $p < 0.01$). To determine the critical geometrical feature for individual neurons, we presented various hand-held three-dimensional objects (such as faces, hands, stuffed animals, plastic fruits and vegetables), and two-dimensional images on a television monitor and as paper mounts. The image of the most effective object was then simplified progressively, and the simplest visual features eliciting the maximal responses were determined[4,5].

Acknowledgements

We thank K.S. Rockland, R. Kado, K. Tanaka, and S. Edelman for comments on the manuscript, M. Fukuda for technical assistance throughout the experiments, and M. Matsumoto for modifying the data acquisition program. This project was partly supported by Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Corporation.

RECEIVED 1 MAY; ACCEPTED 19 JUNE 2001

1. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. Visual properties of neurons in inferotemporal cortex of the Macaque. *J. Neurophysiol.* **35**, 96–111 (1972).
2. Perrett, D. I., Rolls, E. T. & Caan, W. Visual neurones responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* **47**, 329–342 (1982).
3. Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* **4**, 2051–2062 (1984).
4. Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**, 170–189 (1991).
5. Kobatake, E. & Tanaka, K. Neuronal selectivities to complex object features in the ventral pathway of the macaque cerebral cortex. *J. Neurophysiol.* **71**, 856–867 (1994).
6. Ts'o, D. Y., Frostig, R. D., Lieke, E. E. & Grinvald, A. Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science* **249**, 417–420 (1990).
7. Blasdel, G. G. Orientation selectivity, preference, and continuity in monkey striate cortex. *J. Neurosci.* **12**, 3139–3161 (1992).
8. Das, A. & Gilbert, C. D. Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. *Nature* **375**, 780–784 (1995).
9. Roe, A. W. & Ts'o, D. Y. Visual topography in primate V2: multiple representation across functional stripes. *J. Neurosci.* **15**, 3689–3715 (1995).
10. Ghose, G. M. & Ts'o, D. Y. Form processing modules in primate area V4. *J. Neurophysiol.* **77**, 2191–2196 (1997).
11. Malonek, D., Tootell, R. B. & Grinvald, A. Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT. *Proc. R. Soc. Lond. B Biol. Sci.* **258**, 109–119 (1994).
12. Wang, G., Tanaka, K. & Tanifuji, M. Optical imaging of functional organization in the monkey inferotemporal cortex. *Science* **272**, 1665–1668 (1996).
13. Wang, G., Tanifuji, M. & Tanaka, K. Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neurosci. Res.* **32**, 33–46 (1998).
14. Uchida, N., Takahashi, Y. K., Tanifuji, M. & Mori, K. Odor maps in the mammalian olfactory bulb: domain organization and odorant structural features. *Nat. Neurosci.* **3**, 1035–1043 (2000).
15. Gochin, P. M., Miller, E. K., Gross, C. G. & Gerstein, G. L. Functional interactions among neurons in inferior temporal cortex of the awake macaque. *Exp. Brain. Res.* **84**, 505–516 (1991).
16. Fujita, I., Tanaka, K., Ito, M. & Cheng, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* **360**, 343–346 (1992).
17. Marr, D. & Nishihara, H. K. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. B Biol. Sci.* **200**, 269–294 (1978).
18. Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* **94**, 115–147 (1987).
19. Ullman, S. Computation of pattern invariance in brain-like structures. *Neural Net.* **12**, 1021–1036 (1999).
20. Edelman, S., Grill-Spector, K., Kushnir, T. & Malach, R. Toward direct visualization of the internal shape representation space by fMRI. *Psychobiology* **26**, 309–321 (1998).
21. Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L. & Haxby, J. V. Distributed representation of objects in the human ventral visual pathway. *Proc. Natl. Acad. Sci. USA* **96**, 9379–9384 (1999).
22. Wang, Y., Fujita, I. & Murayama, Y. Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nat. Neurosci.* **3**, 807–813 (2000).
23. Baylis, G. C., Rolls, E. T. & Leonard, C. M. Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. *Brain Res.* **342**, 91–102 (1985).
24. Yamane, S., Kaji, S. & Kawano, K. What facial features activate face neurons in the inferotemporal cortex of the monkey? *Exp. Brain Res.* **73**, 209–214 (1988).
25. Hasselmo, M. E., Rolls, E. T., Baylis, G. C. & Nalwa, V. Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Exp. Brain Res.* **75**, 417–429 (1989).
26. Sakai, K. & Miyashita, Y. Neural organization for the long-term memory of paired associates. *Nature* **354**, 152–155 (1991).
27. Erickson, C. A. & Desimone, R. Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J. Neurosci.* **19**, 10404–10416 (1999).
28. Shtoyerman, E., Arieli, A., Slovin, H., Vanzetta, I. & Grinvald, A. Long-term optical imaging and spectroscopy reveal mechanisms underlying the intrinsic signal and stability of cortical maps in V1 of behaving monkeys. *J. Neurosci.* **20**, 8111–8121 (2000).