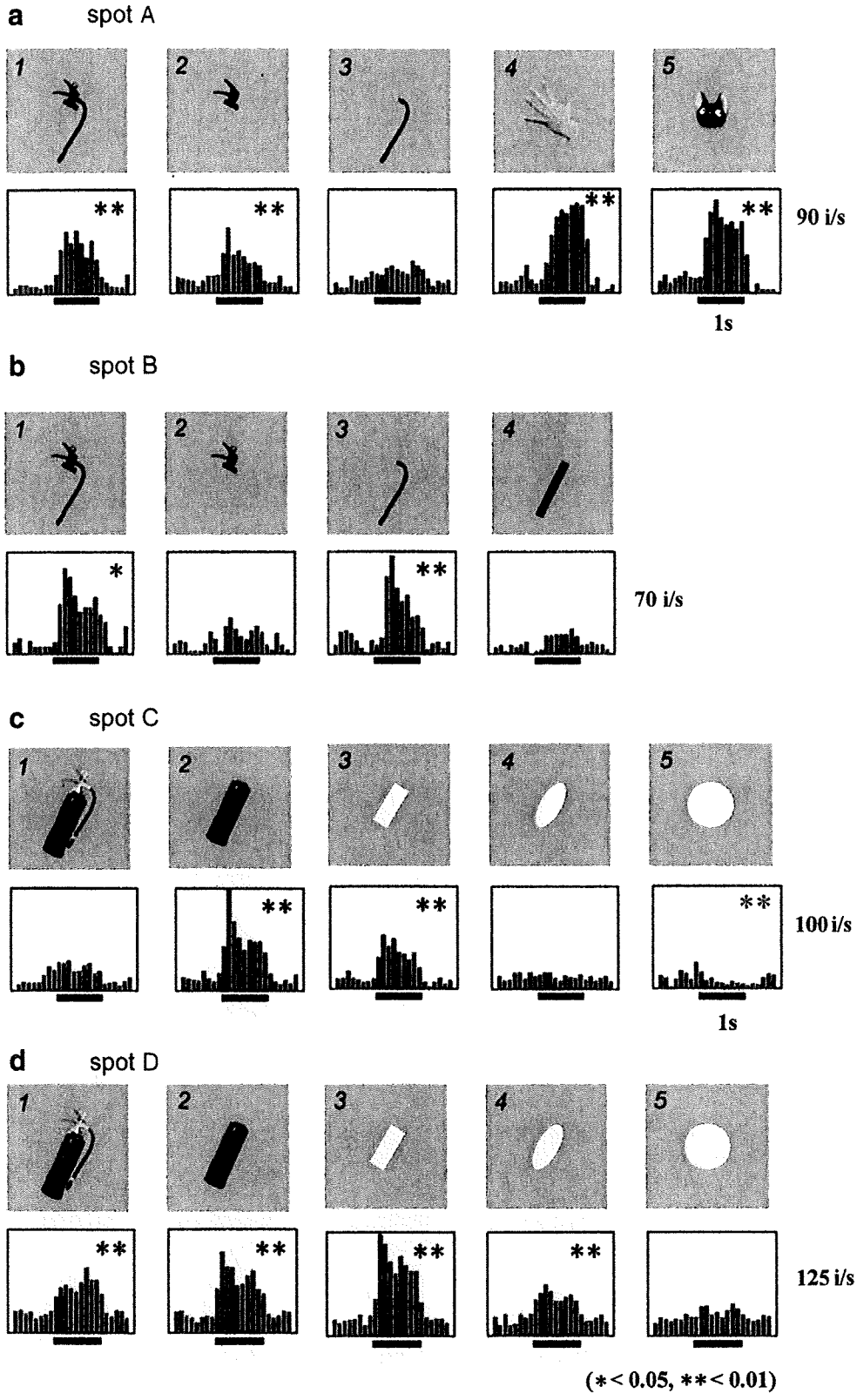

Fig. 5.6 Visual responsiveness of representative cells in spots A–D in Fig. 5.5a. Each histogram (PSTH) shows neuronal responses elicited by the above stimulus. Red asterisks indicate significant inhibition ($p<0.01$). *p <0.05, **$p<0.01$. Scale bars, 1-s periods of stimulation. (Modified from Tsunoda et al. (2001))

OISI revealed that the "cylinder" produced significant responses in both spots C and D (Figs. 5.5a-3), while other stimuli having "the cylinder" as a part (Figs. 5.5a-1, a-2 and a-4) only activated spot D but not spot C. The neural responses of cells in these spots were consistent with the imaging results: the cells in spot C was activated by the "cylinder" but not by the original "fire extinguisher" (Figs. 5.6c-1 and c-2), and a cell in spot D was significantly activated by both stimuli (Fig. 5.6d-1, d-2). The feature critical for the cells in spot D was a "rectangular shape" (Fig. 5.6d-3), but the cells also responded significantly to an "ellipse" (Fig. 5.6d-4). Since there were no responses to a "circle" (Fig. 5.6d-5), an "elongated structure" seemed to be necessary for activation. The simplest visual feature for the cells in spot C was also a "rectangular shape" (Fig. 5.6c-3). In contrast to the cells in spot D, however, none of the cells was activated by an "ellipse" (Fig. 5.6c-4). In addition, all of these cells were inhibited by the "circle" (Fig. 5.6c-5). These results suggest that the response properties of the cells in spot C (Fig. 5.5a) 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," where the hose having a circular shape was attached with the cylinder having a rectangular shape. In general, these results suggest that some of the spots representing a particular feature were inactive when other features were presented together with the feature. This could explain that active spots appeared following simplification of stimuli in some cases.

So far, we have revealed two factors involved in representation of object images in the area TE. First, some of visual features represented by activity spots are local features such as protrusion and asymmetric arc. We refer "local features" to as visual features that occupy part of an object image and are distinguishable from other parts of an object image by their particular shapes, colors, or textures. Second, specific object representation is not made by combinations only of active spots but also of inactive spots. In the case of representation of original fire extinguisher, activation of spots A and B represents the presence of protrusion and curvature in the object, respectively. Activation of spot D indicates that entire structure is elongated, but no activation in spot C further indicates the structure has to be relatively elliptic. In such a way, combinations of inactive as well as active columns help representation of object images to be more specific.

5.5 Representation of Configurational Information Appeared in Object Images

Specific representation of object images by combinations of local features requires mechanisms to represent information about the spatial arrangement of "local features" or about spatial arrangement of parts including local features ("configurational information"). In the previous section, we showed that spot C in Fig. 5.5a was not activated when the hose is attached to the side of the cyl-

inder and makes the entire shape elliptical. This spot, however, would be activated if the hose were secured above the handle where the rectangular shape of cylinder was exposed. From the viewpoint of representing "configurational information," we may consider that activity in spot C carries information about the position of the hose relative to the cylinder, although the way to represent "configurational information" about the relative position of hose and cylinder is indirect.

We searched for the neural substrates that explicitly represent "configurational information." Particularly, we explored representation of a particular spatial relationship, that is, "on top of." This is a typical spatial relationship appeared in object images. For example, the head is above the body, the lampshade is on top of the base, and a pineapple is separated into upper part leaves and lower part fruit. To find candidate spots related to representation of the spatial relationship, we searched for the spots that showed a specific pattern of activation: activation by the object consisting of two parts and the same object with a gap introduced between parts of the object, but no activation by a part alone (Yamane et al. 2006). Among four hemispheres, we found three spots showing this specific pattern of activation (Fig. 5.7). These spots should not simply represent local features of an upper or lower part of the objects because either part is not essential for activation. Moreover, activation by the stimulus with an introduced gap indicates that local features appearing at the junction of two parts, such as sharp connecting corners, are also not essential. Thus, these spots were good candidates that could represent spatial relationship, "on top of." For further characterization of these spots, we made single cellular recordings from these spots, and found that three unique response properties. First, the cells in these spots preferred object images consisting of vertically aligned two parts with a small number of exception (Fig. 5.8). The critical features for these cells determined by the stimulus simplification procedure were also the combinations of vertically aligned two parts except two (Fig. 5.9b-7, b-8). Second, these cells were less sensitive to color, texture, and local shapes of either part. Thus, there were no changes in the responses after removing color and texture during the stimulus simplification procedure (Figs. 5.3 and 5.9). The changes in shapes of the parts did not significantly alter responses of these cells. For example, a neuron, whose critical feature was determined as a combination of a circle and a rectangle, was also significantly activated by a combination of a circle and an ellipse. Third, these cells were highly selective to a particular spatial arrangement of the upper part and the lower part (Fig. 5.10). Most of the cells were maximally activated when the upper and lower parts were vertically aligned or tilted only for 45° from vertical arrangement. We cannot explain the selectivity to a particular spatial alignment of upper and lower parts by changes in retinotopic positions of the upper parts that occurred incidentally during the spatial rearrangements of the parts because the receptive field of these cells covered even larger area in the visual field.

These response properties enable these spots to respond to two-part objects regardless of the local features embedded in either part, but only when the parts are aligned vertically (Fig. 5.8). These results, as well as neurons in spot C in

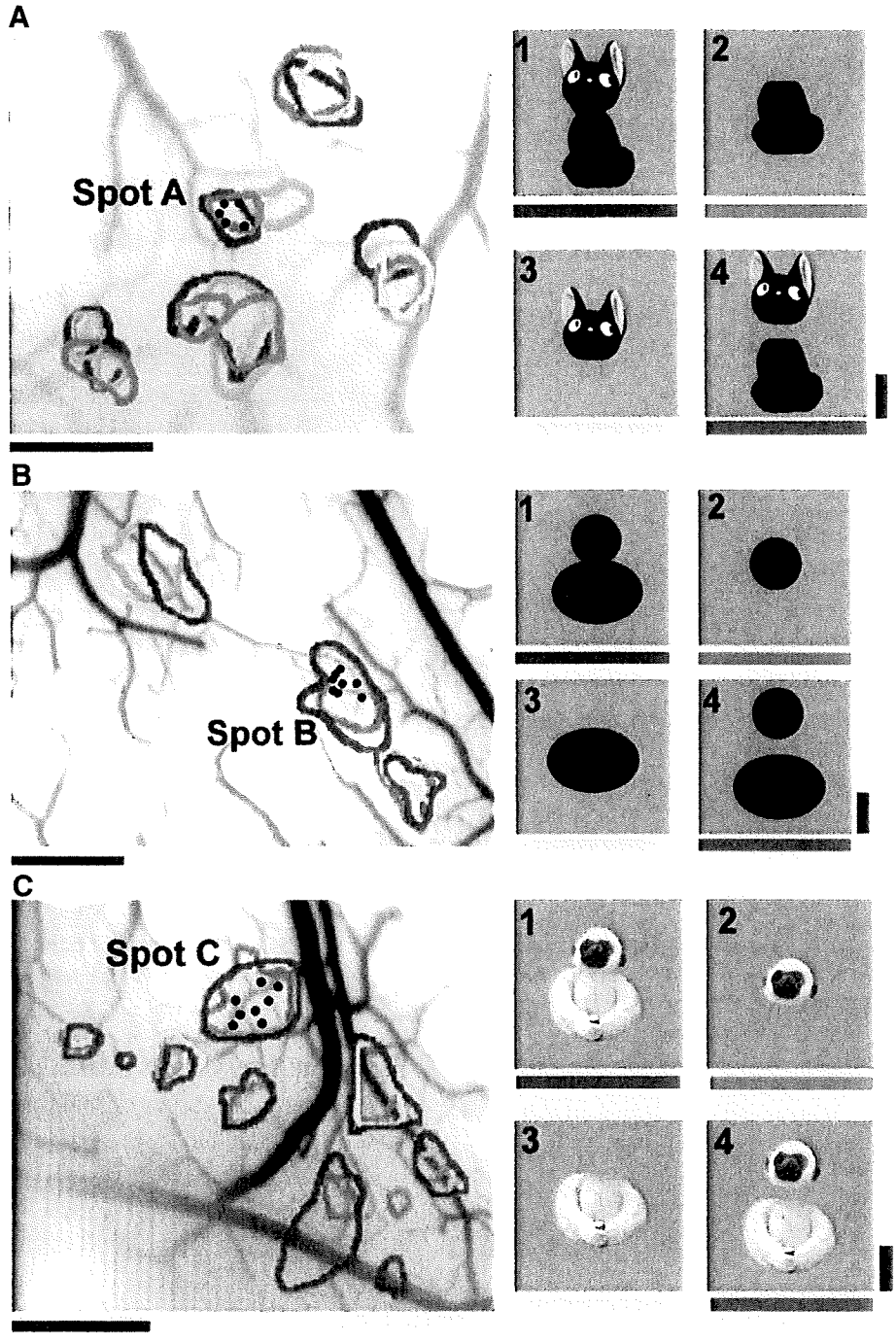

Fig. 5.7 Distribution of activity spots elicited by the stimulus set designed for searching spots related to spatial arrangement between *upper* and *lower* parts. Three spots A, B, and C were identified. *Black dots*, electrode penetration sites (Modified from Yamane et al. (2006))

Fig. 5.5a, suggest that neurons in area TE do not necessarily represent local features but also "configurational information" of the object images. As shown below, we consider that face neurons also represent configurational information

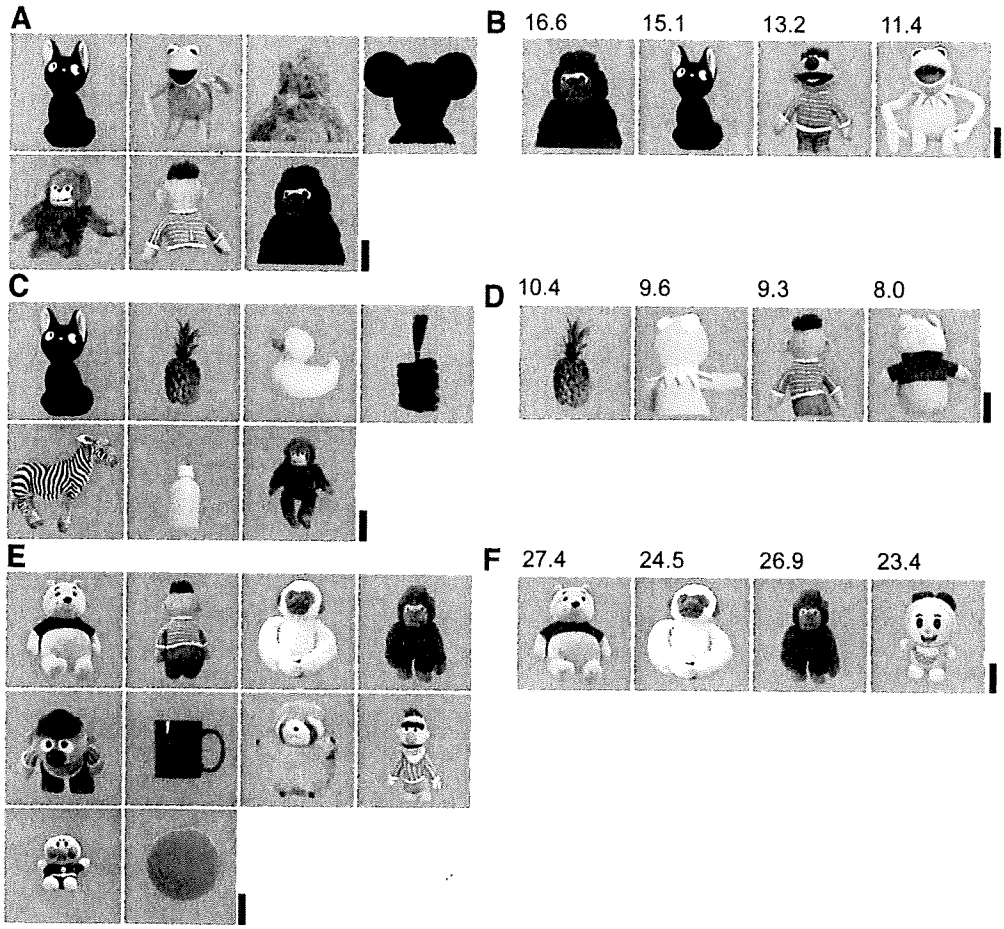

Fig. 5.8 Object stimuli that elicited significant responses for the cells in spots A, B, and C in Fig. 5.7. (**A, C, E**) Object images that elicited the strongest responses out of the 96 objects for cells recorded in spots A (**A**), B (**C**) and C (**E**). (**B, D, F**) The best 4 stimuli out of 96 objects that elicited significant responses for a representative cell in spots A (**B**), B (**D**), and C (**F**). Evoked responses (spikes/s) are indicated above each stimulus image. Scale bar, 5° (Modified from Yamane et al. (2006))

about facial parts. Object images could be specifically represented by combinations of spots representing "local features" and those representing "configurational information."

In the above study, these cells responded to the configuration where the upper part was above the lower part, but not to the configuration where the upper part is below the lower part (Fig. 5.10). This result means that, in some ways, the cells distinguished the upper part and the lower part of objects. In this respect, the cells were not entirely insensitive to local features. At least, the cells could distinguish upper and lower parts. These cells may be sensitive to difference of area of parts. Alternatively, some combination of a curvature in the upper part and in the lower part may be the critical factor (Brincat and Connor 2004).

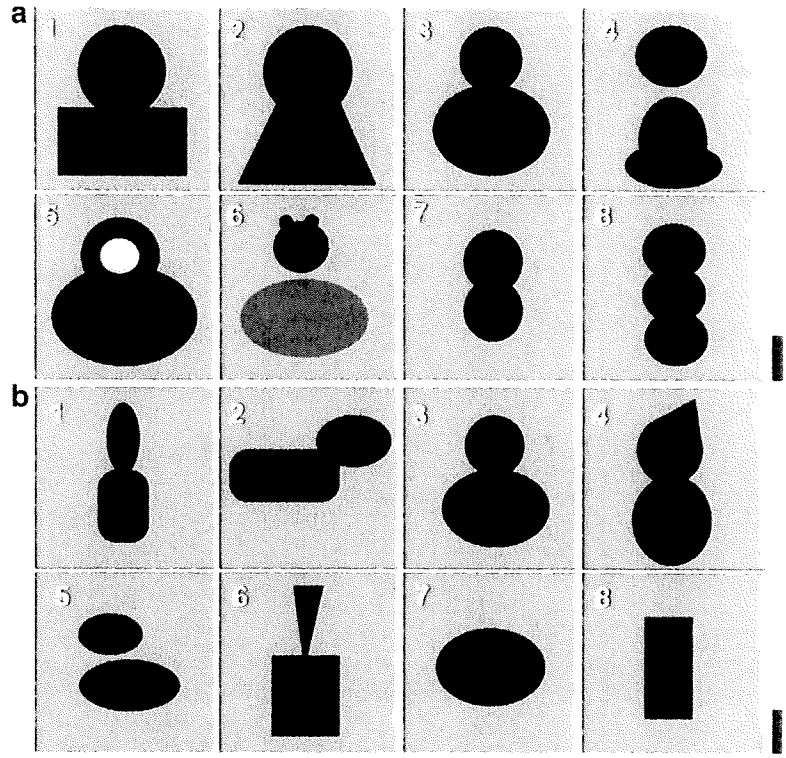

Fig. 5.9 The critical features of the cells in spots A(a) and B(b) related to representation of spatial relationship among parts. The critical features were not investigated for spot C. (Modified from Yamane et al. (2006))

5.6 Face Neurons in Area TE as Ones that Represent Facial Configuration

"Face neurons" is the neurons that respond to "faces," but these responses cannot be explained by specific responses to a part of "face" (Fig. 5.11). For example, a face without eyes did not activate the cell, but there was no activation by "eyes" alone (Fig. 5.11b right most, c). Furthermore, previous studies have shown that "face" with scrambled facial parts do not activate these neurons (Bruce et al. 1981). There are two characteristic properties of "face neurons." First, many of them are tuned to images of faces from a particular vantage point (Fig. 5.11a)(Perrett et al. 1982; Perrett et al. 1991). Second, these cells are less sensitive to difference of individual faces (Perrett et al. 1984; Baylis et al. 1985; Yamane et al. 1988; Young and Yamane 1992). These response properties suggest that the face neurons represent not specific faces but facial configuration.

Intrinsic signal imaging showed that there are spots specifically activated by faces (Fig. 5.12) (Wang et al. 1996; Wang et al. 1998). Thus, face neurons, as well as neurons specifically responding to visual features, are clustered together. Furthermore, activation patterns produced by images of faces from different vantage points revealed that the peaks of activity spots shift along the cortical surface as the

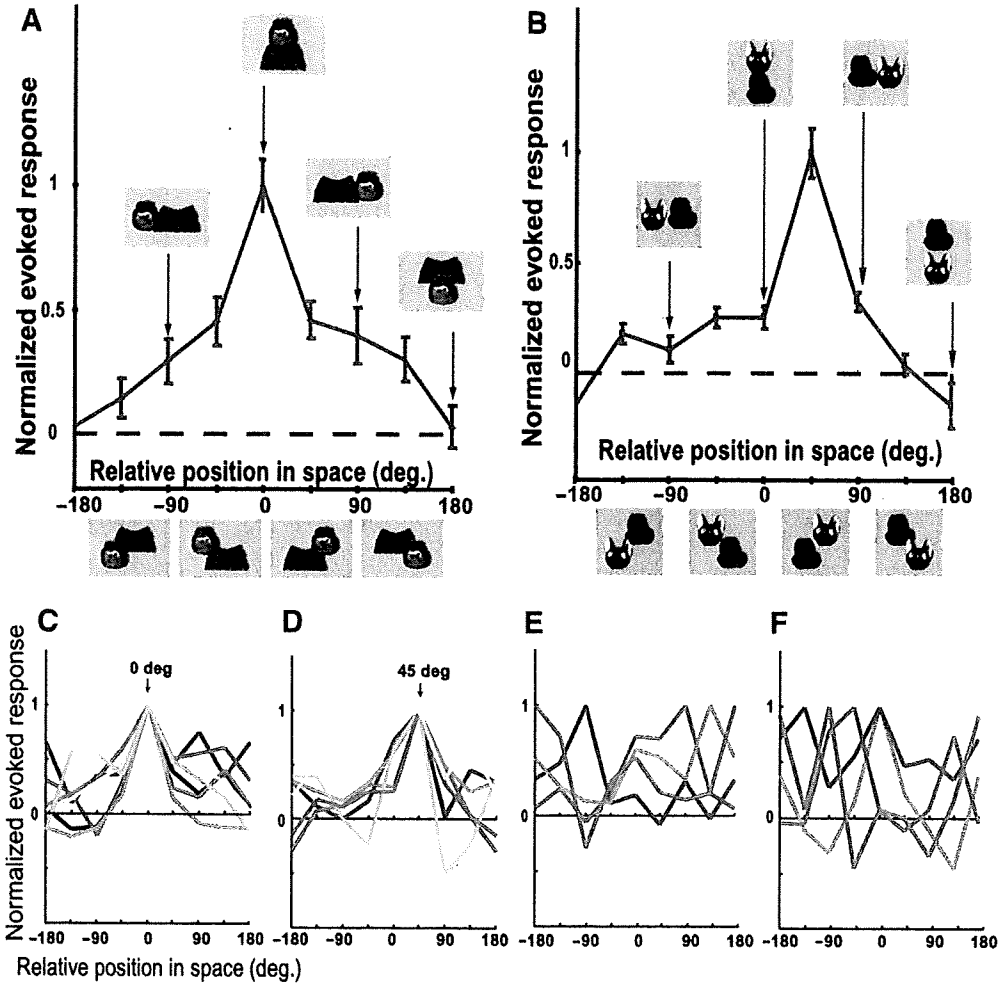

Fig. 5.10 Selectivity of cells to different spatial arrangements of the *upper* and the *lower* parts of object images. The normalized evoked responses (vertical axis) were plotted against the difference in the spatial arrangement of the parts (horizontal axis). The difference in spatial arrangement is defined by the angle between a line connecting the centers of the two parts of the best object stimuli and that of each rearranged stimulus. The pictures of stimuli corresponding to each angle are shown below the plot and also in the insets. The panels (**A**) and (**B**) show responses of representative cells in spots A and B in Fig. 5.7, respectively. (**C-F**) Tuning curves for other cells in others in spots A, B, and C with a single peak at 0°. (**C**), 45°. (**D**) and other angles (**E**), and tuning curves with multiple peaks (**F**). For simplicity, only the mean values of responses are plotted. (Yamane et al. (2006))

face rotates from the left profile to the right profile through the front face. This representation of faces from different vantage points in close vicinity may be important for view-independent recognition of faces.

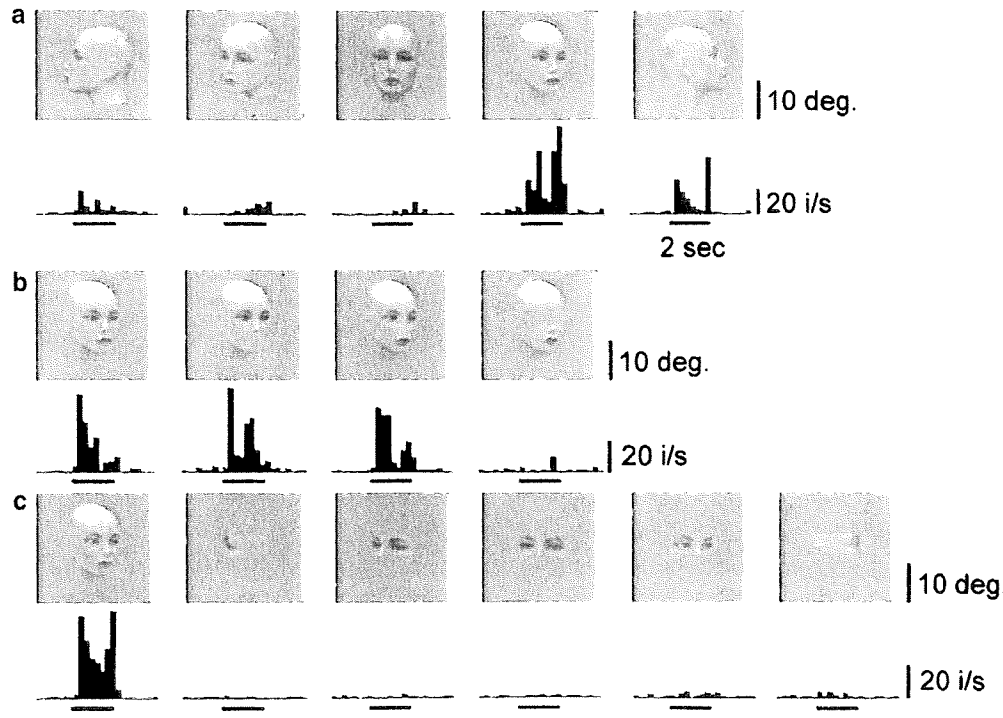

Fig. 5.11 Responses of a face neuron. The recording consists of three sessions. In the first session (**a**), the selectivity of the neuron for different views of a face is demonstrated. In the second and third sessions (**b**, **c**), selectivity of the same neuron for a face and facial parts were examined. Please note that no activation by either a face without eyes (**b**) or an eye alone (**c**) was observed. (Cited from Fukuda and Tanifuji, unpublished observation.)

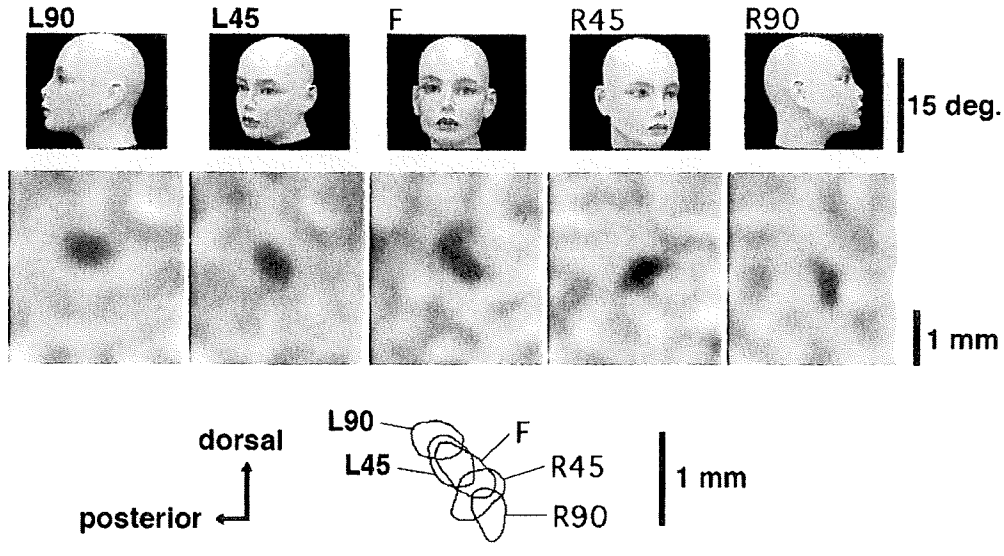

Fig. 5.12 Systematic shift in distribution of activity spots with rotation of the face. Images of the same cortical area (*middle panels*) obtained from five different views of the same mannequin face (*top panels*). The contours of the active spots are superimposed at the *bottom* (Modified from Wang et al. (1996))

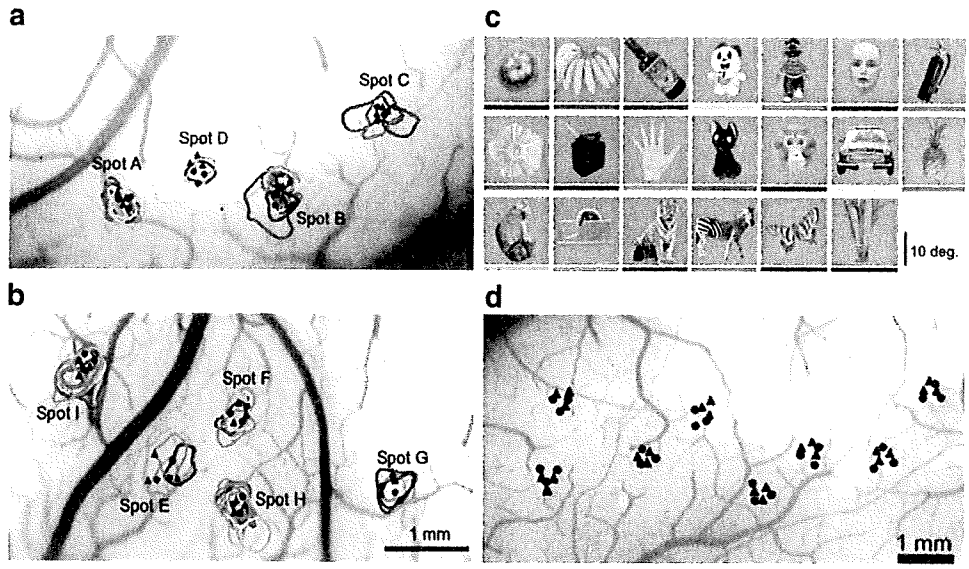

Fig. 5.13 The sites for electrophysiological recordings of single cells and MUs. Electrode penetration sites are given in filled circles and triangles. (**a, b**) The activity spots were shown by colored contours indicating area of activation by object images shown in (**c**). Contours were drawn as in Fig. 5.2. (**c**) Stimulus images for IOSI. The line type and the color are matched to the contours in **a**, and **b**. (**d**) The recording sites in the hemisphere where no IOSI was conducted beforehand. (Modified from Sato et al. (2009))

5.7 Object Representation at Different Levels: Columns and Single Cells Within a Column

Recently, several studies have quantitatively examined object selectivity of IT cells and found that IT cells were tuned to a relatively small number of object images (Tamura et al. 2005; Kreiman et al. 2006; Sato et al. 2009). One measure to quantify specificity of stimulus tuning curves is the sparseness index (Rolls and Tovee 1995). We obtained the value of the sparseness index for TE neurons for randomly chosen 80 object images as 0.19 ± 0.18 ($n=218$). This value, 0.19, means that the neuron responds to only 19% of stimuli, if we approximate neuronal responses to stimuli in an all-or-none fashion (0 or 1). Furthermore, these studies have shown that selectivity of nearby cells was largely different from one another (Tamura et al. 2005; Kreiman et al. 2006; Sato et al. 2009). For example, five object images that evoked the strongest visual responses were different for two neurons even if these neurons were sampled from the same activity spot and spaced only for 150 μm apart (Fig. 5.13, 5.14). The correlation of object selectivity of nearby two cells within the same activity spots was as low as 0.15 in the mean value of the correlation coefficient, and the number of cell pairs that had significant correlation was only 28.5% ($p<0.05$) (Fig. 5.15a) (Sato et al. 2009). The difference in object selectivity of nearby cells was not an artifact due to trial-by-trial variability of neuronal responses. The correlation coefficient expected from the trial-by-trial variability

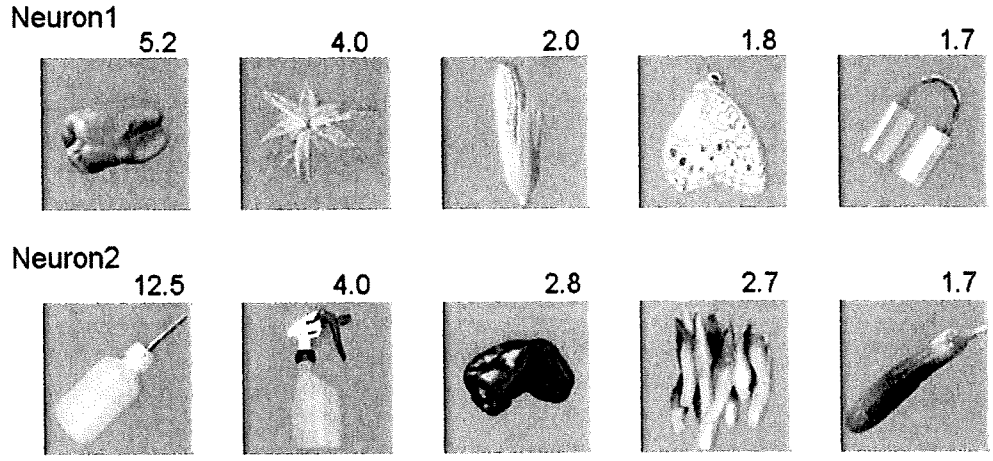

Fig. 5.14 The best five stimuli of two nearby cells spaced 150 μm apart. These two cells were isolated from spot A in Fig. 5.13. The number above each stimulus was the evoked response by the stimulus (spikes/s). The correlation coefficient of selectivity of these two cells for 100 object images was 0.07

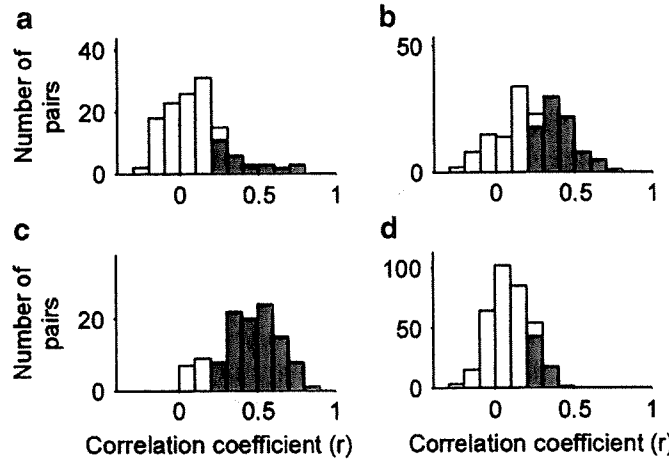

Fig. 5.15 Similarity in object selectivity between pairs quantified by correlation coefficient of evoked responses to 80 object stimuli. The pairs are made of single isolated cells (**a**), and MUs (**b**). Proportions of pairs that had significant correlation were 28.5% (**a**) and 60.0% (**b**). In (**c**), the correlation coefficient was calculated for the pair of MUs and the avgMU obtained from the same columns. Proportion of MUs that had significant correlation in object selectivity with avgMUs was 85.2%. On the other hand, in (**d**), we calculated correlation coefficient for the piars of MUs and avgMU obtained from different columns, Proportion of MUs that had significant correlation was only 16.4%. Thus, the common property represented by the avgMU was unique to the activity spot where the avgMU was calculated and was different from other spots. Horizontal axis, the value of correlation coefficient. The columns with red color represent group of pairs that has statistically significant correlation ($p < 0.05$). (Modified from Sato et al. (2009))

was around 0.4 in correlation coefficient, and this value is higher than the value of similarity in object selectivity of nearby isolated cells. Thus, although evidence suggests a columnar organization in IT cortex (Fujita et al. 1992), object selectivity of nearby cells seemingly contradicts to the column hypothesis in IT cortex.

Furthermore, these findings raise a possibility that object images are specifically represented even by combinations of single cells in a local region that have different response property.

To resolve this contradiction, we analysed object responses of single cells and multi-units (MUs) recorded from activity spots (Fig. 5.13), and explained the reason for the difference in object selectivity of nearby cells that cell-specific response property obscured the common response property across the cells within a columnar region (Sato et al. 2009). In accordance with this explanation, the number of pairs of multiunits (MUs) (60%) showing significant correlation in object selectivity was higher than that of single neuron pairs (28.5%) probably because the variation caused by cell-specific responses were averaged across the cells and the common property became apparent in MU responses (Fig. 5.15a, b). Furthermore, we found that this common property (quantified as the average of MU activities recorded from the same activity spot (avg MU activity)) was different from activity spot to activity spot (Fig. 5.15c, d). Thus, although there is cell-to-cell variability in object responses, the columnar structure is maintained since these cells share a common response property. The relationship between the common property and cell-specific property is demonstrated in Fig. 5.16 by plotting object responses of individual cells against descending order of object responses of the avgMU activity. Furthermore, the PCA analysis of MUs recorded from multiple activity spots revealed that one activity spot was characterized by one or a few common properties in object responses (Fig. 5.17). As we mentioned earlier, analysis of the critical features of individual cells first revealed the columnar organization in area TE (Fujita et al. 1992). We consider that the procedure of stimulus simplification made them find the visual feature that represents the common property across the cells within a columnar region, although the authors of the paper did not intended to search for the common property across the cells. Recently, it has been shown that nearby cells in V1 also respond differently to natural images although these cells should have similar orientation preference (Yen et al. 2007). Thus, the differentiation between the response property specific at a single cell level and the response property at the columnar level may be a universal characteristic in cerebral cortices.

The critical question particularly in area TE is how to relate object representation by combinations of columns and a possible representation by combinations of single cells in a local region. Taking into account trial-by-trial variation of neuronal activity, we consider that representation at the columnar level has primary importance. Neuronal firing approximately follows Poisson distribution, and thus, trial-by-trial variability in neuronal responses is higher for the good stimuli that evoked strong mean responses than the stimuli that evoked weak mean responses. Thus, in general, it is difficult to extract information about the stimulus from activity in a single trial of a single neuron. Since we recognize an object instantaneously without trial-by-trial averaging, one plausible way to extract reliable sensory information is to take an ensemble average of responses of nearby neurons. Our results above point out that this ensemble average reflects tuning specificity of feature columns, and the specificity is generally different from single cells. Accordingly, we consider that recognition of object images presented at a

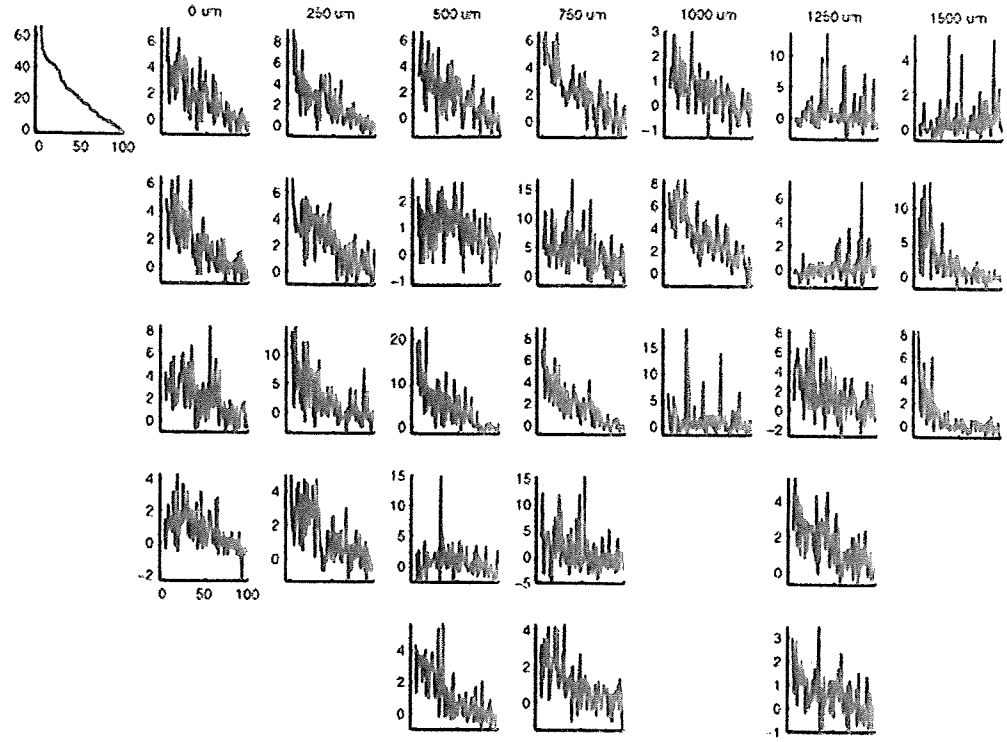

Fig. 5.16 The common response property revealed in tuning curves of the individual cells. Evoked responses of each cells were plotted against object stimuli rank ordered by the avgMU responses. The graph at the *left upper* corner represents the tuning curve of averaged MUs and the rests represent tuning curves of single cells at different depths. Depth of cells in each column is indicated at the *top*. Horizontal axes are rank-ordered according to the magnitude of evoked responses of averaged MUs to the 100 object stimuli in descending order. Vertical axes represent mean firing rate (spikes/s). As you see in single cell tuning curves, though the object that elicited the strongest responses was different from cell to cell, tuning curves of these cells have general tendency that rightward objects evoked weaker responses than leftward objects. (Modified from Sato et al. (2009))

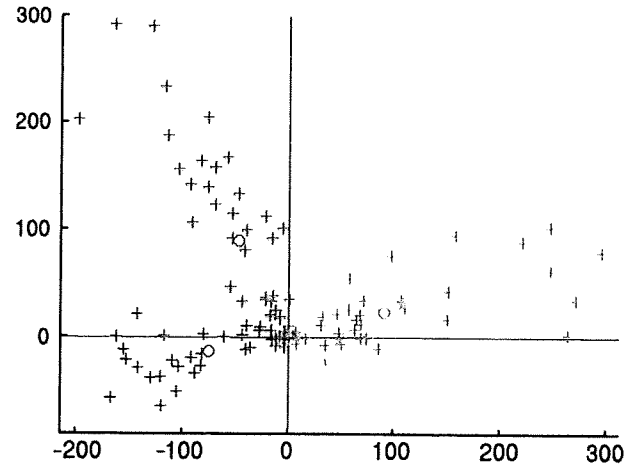

Fig. 5.17 One response property for each spots is approximated in comparison with difference in response property across activity spots. Object Responses of MUs of three spots (crosses; different color indicates different spots) are plotted in a multiple dimensional space made of 80 object images, and projection onto the 2-dimensional plane that includes responses of avgMU in three activity spots (open circles) is shown. (Modified from Sato et al. (2009))

glance is achieved based on representation of object images not at the level of single cells but at the columnar level. It is still an open question in what occasion representation of object images at the level of single cells is useful (but see Summery and Discussion).

Finally, unlike the columnar organizations in V1, we propose that the columnar organization may not cover entire cortex uniformly (Sato et al. 2009). In one hemisphere, we analyzed similarity in object selectivity of cells not from activity spots identified by IOSI but from arbitrary assigned local regions that have about the size of activity spots (Fig. 5.13d). If the columnar organization uniformly covers the cortex, the results would be the same as those from activity spots. However, we found that correlation in object selectivity among MU pairs greatly reduced. This observation suggests that in part of the cortex common property across the cells are not obvious than that in the activity spots.

5.8 Summary and Discussion

In this chapter, we have described object representation in IT cortex revealed by OISI: object images were represented by combinations of active and inactive feature columns, and these feature columns carry information about both local visual features as well as global features such as configuration of object parts. Though we showed representation of these two types of features in separate experiments, we do not intend to emphasize dichotomy between local and global features. It could be the case that some columns represent a visual feature that carries partly global and partly local information although such columns have not been explored yet. The visual features shown here, that is, protrusions, curvature, and rectangular shape for local features and the vertical alignment of parts and facial configuration for global features, are representatives that are easily described with a simple concept. As mentioned in an earlier part of the chapter, plasticity of IT cells make them adjust their response property to the new environment that monkeys react to. Thus, IT cells could tune to the visual features that are useful in certain behavior regardless of whether the represented features are local, global, or include both. Thus, in general, we may not easily describe visual features represented in IT cortex. For example, it would be difficult to characterize the visual features useful for Japanese to recognize Kanji characters at a glance and for Arabic to recognize Arabic characters at a glance.

Because of trial-by-trial variability in neuronal responses, we have discussed that object representation at the level of columns may have primary importance. Then, why do individual cells have a relatively high object tuning property that is different from cell to cell? One thought would be that variability in object responses among cells is necessary for shaping a response property essential for object representation at the columnar level. Alternatively, however, the object representation at the level of single cells may play an essential role in object recognition under certain circumstances. For example, when we inspect an object image, we gaze at the same portion of the image multiple times (Yarbus 1967).

Repeated fixations may play a role in making trial-by-trial averaging of responses at the level of single cells. Then, object recognition may have two stages: the first stage is to capture an object image at the level of feature columns by using ensemble averaging, and for the following careful inspection of the object image, we use representation at the level of single cells with trial-by-trial averaging through repeated fixations. At present, there is no evidence supporting this two stage model. At least, we need to investigate fundamental problems related to solving trial-by-trial variation in neuronal responses. In the case of ensemble averaging, it is essential whether the trail-by-trail variations are uncorrelated across neurons or not. On the other hand, in the case of solving trial-by-trial variations with repeated fixations, a certain kind of short term memory systems is required for holding sensory responses obtained in different time.

As described above, columnar organizations have a potential role in instantaneously and reliably detecting visual features at the columnar level. Thus, to maximize efficiency of plastic changes in IT cortex, plasticity may not only shape response property of individual cells depending on experience but also it may make clusters of these cells and generate new feature columns that are adjusted to the new environment. As described above, we suggest that columnar organization for visual features do not cover the entire region of IT, the rest of the region would be reserved area for generating new feature columns where neurons represent a common visual feature that is useful in certain behavior.

Since the results shown here were obtained from anesthetized monkeys, investigations of dynamic properties of object representation with behaving monkeys are required to relate object representation to recognition. For example, we have examined representation of a single object presented in the visual field. This experimental condition is very much far from natural conditions where many objects are presented simultaneously. Thus, one unsolved problem is how to deal with multiple objects in the visual field. One possible mechanism is proposed by Reynolds et al. (1999). They found that, regardless of whether the visual stimulus was presented in isolation or together with the other stimulus, visual responses of single V4 cell were the same when monkey pays attention to the stimulus. If this is the case, when multiple objects were presented to a monkey, spatial patterns of activity would change from one pattern to another depending on which object the monkey pays attention to. To address this kind of questions, investigations on representation by combinations of feature columns in behaving animals are required. So far in the literature, most of the investigations with behaving monkeys are based on physiological recordings from a single neuron at a time. Since the results presented here revealed that object representation is made of combinations of feature columns, we think that recordings from population activity particularly at the columnar level is essential. However, because of the slow time course of intrinsic signals (Fig. 5.1), OISI may not be the appropriate technique to address object recognition with behaving monkeys in which representation of object images supposed to be updated in a range of a few hundred milliseconds. Novel techniques with high temporal resolution as well as spatial resolution are required. Recently developed densely arranged multiple electrode arrays would be one of the such techniques (deCharms et al. 1999; Miyakawa et al. 2007).

References

Baylis GC, Rolls ET, Leonard CM (1985) 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(1):91–102

Brincat SL, Connor CE (2004) Underlying principles of visual shape selectivity in posterior inferotemporal cortex. Nat Neurosci 7:880–886

Bruce C, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 46(2):369–384

Das A, Gilbert CD (1995) Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. Nature 375:780–784

deCharms RC, Blake DT, Merzenich MM (1999) A multielectrode implant device for the cerebral cortex. J Neurosci Methods 93:27–35

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci 4(8):2051–2062

Fujita I, Tanaka K, Ito M, Cheng K (1992) Columns for visual features of objects in monkey inferotemporal cortex. Nature 360(6402):343–346

Fukuda M, Rajagopalan UM, Homma R, Matsumoto M, Nishizaki M, Tanifuji M (2005) Localization of activity-dependent changes in blood volume to submilimeter-scale functional domains in cat visual cortex. Cereb Cortex 15(6):823–833

Grinvald A, Shoham D, Shmuel A, Glaser D, Vanzetta I, Shtoyerman E, Slovin H, Wijnbergen C, Hildesheim R, Arieli A (1999) In-vivo optical imaging of cortical architecture and dynamics. In: Windhorst U and Johansson H (eds) Modern techniques in neuroscience research. Springer, Berlin Heidelberg New York pp. 893–970

Gross CG (1994) How inferior temporal cortex became a visual area. Cereb Cortex 5:455–469

Gross CG, Bender DB, Gerstein GL (1979) Activity of inferior temporal neurons in behaving monkeys. Neuropsychology 17:215–229

Homma R, Tanifuji M (2003) Comparison of functional MAPs in macaque area TE revealed by *in vivo* optical imaging with voltage-sensitive dye and intrinsic signal imaging. Abstr. viewer/Itinerary Planner. Society for Neuroscience, Washington, DC. 818.21

Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J Neurophysiol 71(3):856–867

Kobatake E, Wang G, Tanaka K (1998) Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. J Neurophysiol 80(1):324–330

Kreiman G, Hung CP, Kraskov A, Quiroga RQ, Poggio T, DiCarlo JJ (2006) Object selectivity of local field potentials and spikes in the macaque inferior temporal cortex. Neuron 49:433–445

MacVicar BA, Hockman D (1991) Imaging of synaptically evoked intrinsic optical signals in hippocampal slices. J Neurosci 11:1458–1469

Miyakawa N, Vidal-Naquet M, Blake D, Merzenich M, Tanifuji M (2007). Activities from combination of columns in macaque area TE can encode object identity across viewing angles, Abstr. viewer/Itinerary Planner. Society for Neuroscience, San Diego. Online, 554.11.

Perrett DI, Rolls ET, Caan W (1982) Visual neurones responsive to faces in the monkey temporal cortex. Exp Brain Res 47(3):329–342

Perrett DI, Oram MW, Harries MH, Bevan R, Hietanen JK, Benson PJ, Thomas S (1991) Viewer-centred and object-centred coding of heads in the macaque temporal cortex. Exp Brain Res 86(1):159–173

Perrett DI, Smith PA, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves MA (1984) Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol 3(4):197–208

Reynolds JH, Chelazzi L, Desimone R (1999) Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci 19:1736–1753

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

Sato T, Uchida G, Tanifuji M (2009) Cortical columnar organization is reconsidered in intorial tempral cortex. Coreb Corlex 19:1870–1880

Tamura H, Kaneko H, Fujita I (2005) Quantitative analysis of functional clustering of neurons in the macaque inferior temporal cortex. Neurosci Res 52:311–322

Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol 66(1):170–189

Tsunoda K, Oguchi Y, Hanazono G, Tanifuji M (2004) Mapping cone- and rod-induced retinal responsiveness in macaque retina by optical imaging. Invest Ophthalmol Vis Sci 45(10):3820–3826

Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M (2001) Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat Neurosci 4(8):832–838

Vanzetta I, Grinvald A (1999) Increased cortical oxidative metabolism due to sensory stimulation: implications for functional brain imaging. Science 286(5444):1555–1558

Vanzetta I, Solvin H, Omer DB, Grinvald A (2004) Columnar resolution of blood volume and oximetry functional maps in the behaving monkey; implications for fMRI. Neuron 42(5):843–854

Wang G, Tanaka K, Tanifuji M (1996) Optical imaging of functional organization in the monkey inferotemporal cortex. Science 272(5268):1665–1668

Wang G, Tanifuji M, Tanaka K (1998) Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. Neurosci Res 32(1):33–46

Yarbus AL (1967) Eye movements and vision. Plenum Press, New York

Yamane S, Kaji S, Kawano K (1988) What facial features activate face neurons in the inferotemporal cortex of the monkey? Exp Brain Res 73(1):209–214

Yamane Y, Tsunoda K, Matsumoto M, Phillips AN, Tanifuji M (2006) Representation of the spatial relationship among object parts by neruons in macaque inferotemporal cortex. J Neurophysiol 96:3147–3156

Yen S-C, Baker J, Gray CM (2007) Heterogeneity in the Responses of adjacent neurons to natural stimuli in cat striate cortex. J Neurophysiol 97:1326–1341

Young MP, Yamane S (1992) Sparse population coding of faces in the inferotemporal cortex. Science 256(5061):1327–1331

Novel Snapshot Imaging of Photoreceptor Bleaching in Macaque and Human Retinas

Yoko Kazato,[1,2,3] Naohisa Shibata,[4] Gen Hanazono,[1,2,5] Wataru Suzuki,[1,2] Manabu Tanifuji,[2] and Kazushige Tsunoda[1,2]

[1]Laboratory of Visual Physiology, National Institute of Sensory Organs, Tokyo Japan

[2]Laboratory for Integrative Neural Systems, Brain Science Institute, Riken, Saitama, Japan

[3]Department of Ophthalmology, Nihon University School of Medicine, Tokyo Japan

[4]Nidek Co., LTD., Aichi, Japan

[5]Department of Ophthalmology, Kikkoman General Hospital, Chiba, Japan

Running title: Snapshot Imaging of Photoreceptor Bleaching

Supported by a grant for Research in Sensory and Communicative Disorders' from the Ministry of Health, Labor and Welfare, Japan

Corresponding author: Kazushige Tsunoda, Laboratory of Visual Physiology, National Institute of Sensory Organs, Tokyo Japan, 2-5-1 Higashigaoka, Meguroku, Tokyo 1528902, Japan;
tsunodakazushige@kankakuki.go.jp
Tel: +81-3-3411-0111
Fax: +81-3-3411-0185

Abstract

Purpose: Various methods were used to obtain a topographic map of the bleached photopigments in human retinas in the past. The purpose of this study was to determine whether the bleaching topography of the photoreceptors can be obtained by snapshot imaging reflectometry.

Methods: Four to five fundus photographs were taken by white flashes at intervals of 4 s from one rhesus monkey and three healthy human subjects, with a commercial fundus camera with minimal modifications. The flash-induced reflectance increases (bleaching) were calculated by dividing the reflectance of the first image into the subsequent images pixel-by-pixel.

Results: The topography of the bleached macula corresponded well with the anatomical distribution of the cones. The ratio of reflectance changes in the center to that in the surrounding tissue was high for the red and low for the green and blue images. These results indicate that the reflectivity changes were not artifacts but were derived from changes in the photopigment density in the cones and rods.

Conclusions: The topography of bleached photoreceptors obtained with a commercial fundus camera from one monkey and three healthy human subjects demonstrate that this technique has a potential of being a new clinical method to examine photoreceptor functions both in normal and in diseased retinas.

Keywords: bleaching, densitometry, fundus camera, imaging reflectometry, snapshot photography

Introduction

The morphology of the retina can be imaged with very high spatial resolution by recent imaging techniques, e.g., scanning laser ophthalmoscopy (SLO)[1, 2] and optical coherent tomography (OCT).[3, 4] However, these imaging instruments do not allow an objective assessment of retinal functions, which is essential to making correct diagnoses of various retinal disorders. Because the quality of vision, such as visual acuity and visual fields, depends not only on the anatomical density of the photoreceptors but also on the neural function of the photoreceptors, mapping the retinal responsiveness is important both from the physiological and the clinical point of view.

At present, the functional topography of cone-derived retinal responsiveness is obtained by multifocal electroretinograms (mfERGs) in the clinic,[5] but the spatial resolution of mfERGs is limited. The technique of intrinsic signal imaging with infrared light[6-14] and functional OCT[15-17] have been recently used to map the light scattering changes of the photoreceptors following neuronal activities. However, these techniques have not been brought into the clinic mainly because of their low signal-to-noise ratios (SNRs).

Some fifty years ago, the bleaching of the photopigments was determined quantitatively by measuring the reflectance changes during the bleaching and regeneration processes in human retinas. This method was used to determine the in vivo kinetics of the photopigments in the cone and rod photoreceptors precisely.[18-22] Retinal densitometry was later used to measure the spatial and spectral distribution of the reflectance changes by examining images obtained by either a fundus camera or by a scanning laser ophthalmocope (SLO), i.e., imaging fundus reflectometry.[14, 23-31] These techniques allowed the investigators to objectively and non-invasively map the cones and rods as bleach-derived light reflectance changes both in normal and diseased eyes.

Imaging fundus reflectometry, however, has fundamental problems both theoretically and technically,[31, 32] and has not been wildly used in the clinic in spite of the numerous reports published from different laboratories. One of the practical problems is that the recording period during which the subject must fixate a target steadily is long and uncomfortable. This is a common problem shared by many of the imaging techniques, such as OCT, that are used to study the ocular fundus. However, there is a much more critical problem in reflectometry measurements which is that a small displacement of the fundus image during the long fixation period leads to a loss of important densitometric information.

To overcome this limitation, we have measured the retinal reflectance changes by taking consecutive snapshot images with short duration flashes (<1.0 msec), and analyzing the density of corresponding areas of the retina pixel-by-pixel as a function of time. The topographic maps of the bleached photoreceptors were obtained from three healthy human subjects within 12 s, and the reliability of the results was confirmed in an experiment on an anesthetized monkey. A commercial

fundus camera with minimal modification was used and the bleaching topography could be obtained through simple analyses of the images.

Methods

Apparatus

A slightly modified commercial fundus camera (AFC-210, NIDEK CO., LTD., Japan) was used. The field of view covered 45 degrees of the ocular fundus. To observe the fundus, light from a halogen lamp (fig. 1-B, **I**) passed through an infrared interference filter (**H**) and the optical system of the camera (**F, B and A**) to illuminate the retina, and the retinal image was monitored by the image on a charge-coupled device (CCD) camera. This image was used to adjust the field viewed and to focus on the retina before the recording. A small fixation target of a green light-emitting diode (LED), which was masked by a triangular pattern, was placed in the optical path (**E**). A movable mirror (**C**) was placed between the objective lens and the CCD camera, which was used to change the optical pathway either to the CCD camera for focusing or to a CMOS camera for recording the images (EOS-5D, Canon Inc., Japan). Just before the flash, the mirror was mechanically displaced for 1/50 sec so that the retinal image could be photographed by the CMOS camera. A ring-shaped Xenon strobe (**G**) (flash duration: <1.0 msec) was placed in the optical pathway of the halogen lamp. The variation in the intensity of the Xenon flash strobe was less than $\pm 1.0\%$, negligible considering that the reflectance changes at the fovea was 20-40% in the monkey and 15-20% in humans.

To enable a more accurate fixation in this imaging reflectometric method, the fixation target was modified as follows: the circular mask for the fixation target was replaced by a triangular one, and the subjects were instructed to fixate on one apex of that triangle during the whole recording period. . The intensity of the target illumination was set lower than the threshold level of bleaching in this system.

Data analyses

The raw images were first converted to TIFF images (350 dpi, 16 bit, 4368 x 2912 pixels) using 'Digital Photo Professional, Ver. 2.1.1.4', developed by the manufacturer of the fundus camera. These images have separate information for blue, green and red components of the images which correspond to the three peak sensitivities of the CMOS: 460, 540, and 640 nm (**Fig. 2, raw images**). However, each of the spectral characteristics of the CMOS may not be equally converted to the TIFF image, and unfortunately, the technical details of the color-image processing are not disclosed by the manufacturer. The peak sensitivities of the CMOS were measured by the investigators using a spectral light source (OL490 Agile Light Source, Optronic Laboratories, Inc., FL, USA).

In the human measurements, the alignment of consecutive images was performed by normalized grayscale correlation, to correct for misalignments of the retinal images due to eye movements (x-y axes and rotation) before the differential analyses. When estimating the location of the true correlation peak, the alignment was improved by fitting an analytical model of the correlation peak based on a similarity interpolation estimation, which allowed for sub-pixel estimation. We have applied the

112 algorithm from 'Matrox imaging library, version 9.0' to estimate the exact position of the correlation
113 peak with 0.01 pixel-resolution. If the motion-derived displacements among the consecutive flashes
114 were too large to be corrected by the off-line alignment, the data were discarded. The flash-induced
115 reflectance increases (bleaching) were calculated by dividing the reflectance of the first image by the
116 subsequent images, pixel by pixel (**Fig. 2, Diff. images**). These differential images were spatially
117 smoothed by a simple moving average (32 x 32 pixels in monkey and 54 x 54 in human) to reduce the
118 noise of the CMOS camera before displaying both the profile and the topographic images of
119 bleaching. These differential images were separated into red, green, and blue components (**Fig. 2**).
120
121 *Subjects and recording protocols*
122 Snapshot images were taken from one rhesus monkey (*Macaca mulatta*) and three healthy human
123 subjects (**Subject 1,** 43 y/o man; **Subject 2,** 38 y/o man; and **Subject 3,** 33 y/o woman).
124 The experimental protocol for the monkey was approved by the Experimental Animal Committee
125 of the RIKEN Institute, and all experimental procedures in animals were carried out in accordance
126 with the guidelines of the Riken Institute. For the human experiments, the procedures used conformed
127 to the tenets of the Declaration of Helsinki, and informed consent was obtained from each individual
128 who participated in this study after an explanation of the procedures used. The project was approved
129 by the Institutional Review Board of National Institute of Sensory Organs.
130 Monkeys were first injected intramuscularly with atropine sulphate (0.08 mg/kg) and then
131 anesthetized with droperidol (0.25 mg/kg) and ketamine (5.0 mg/kg). The animals were paralyzed
132 with vecuronium bromide (0.1-0.2 mg/kg/hour) to suppress eye movements, and artificially ventilated
133 with a mixture of 70% N_2O, 30% O_2, and 1.0-1.5% isoflurane. The EEGs, ECGs, expired CO_2, and
134 rectal temperature were monitored continuously throughout the experiments. Before recording, the
135 pupils were fully dilated with topical tropicamide (0.5%) and phenylephrine hydrochloride (0.5%).
136 After 30 minutes of dark-adaptation, the retina was focused on the camera plane by observing the
137 infrared images with the CCD camera. Then, 5 fundus photographs were taken by white flashes with 4
138 s intervals for a total of 16 s. The flash intensity was 101.0 cds/m^2.
139 The retinas of the human subjects were measured under room illumination without
140 dark-adaptation. Following pupillary dilation, each subject's head was stabilized on a head- and
141 chin-rest. During the whole duration of the measurements, the subject fixated on one apex of a
142 triangle which was illuminated by a green LED. Four fundus photographs were taken by white flashes
143 at 4 s intervals. The flash intensity was either 31.8 (**Subject 1**) or 63.6 (**Subject 2 and 3**) cds/m^2.
144
145 **Results**
146 In the monkey, the reflectance of the whole recording region of the retina was increased by the
147 consecutive flashes (**Fig. 3**). The increase in the reflectance in all 3 colors always peaked at the fovea.
148 The bleaching profiles of 4 consecutive images in red showed that the reflectance gradually increased
149 at the fovea (15% in 2nd/1st and 23% in 5th/1st images), whereas the increase was not apparent in the