## Text

A rapid diagnosis of patients with acute visual loss is critical[1], but is difficult if the retina appears normal ophthalmoscopically. We report a patient who presented with an acute unilateral visual decrease and a central scotoma.

### Case Report

A 75-year-old man complained of a sudden and painless decrease of vision in his left eye. He had undergone surgery for an unruptured intracranial aneurysms 20 years earlier, and was taking 7 mg/day of systemic predonisolone for rheumatoid arthritis. He also had diabetic mellitus without retinopathy for 5 years. He was being followed for a left hemianopsia and normal tension glaucoma for 3 years.

On examination, a left relative afferent papillary defect (RAPD) was observed, and the visual acuity (VA) was 20/30 OD and 20/2000 OS. All of the ocular findings were normal except for enlarged disc cupping OU (Figure 1). Fluorescein angiography showed a delay in the arm-to-retina circulation time of 20.0 seconds (Figure 1). Neither retinal emboli nor localized filling delay were observed. Goldmann perimetry showed a left quadrantic homonymous hemianopsia, and a small central scotoma and peripheral constriction, which were new, in the left eye (Figure 2). Focal macular electroretinograms (FMERGs) were recorded two hours after onset as described (see supplemental figure)[2,3].

The FMERGs were decreased in the left eye (Figure 3) indicating that the visual dysfunction was retinal in origin. Three hours later, the left VA improved to 20/250 spontaneously. On the

following day, the VA had improved to 20/30, and the amplitudes of the FMERGs, full-field ERGs, and pattern visual evoked response were normal (Figure 4). The visual field performed three months later showed disappearance of the central scotoma and improvement of isopters (Figure 1). The fundus remained normal. Ultrasound echography revealed no stenosis of the carotid artery. Blood examination revealed rheumatoid factor and high HbA1c but otherwise normal.

We concluded that the acute visual loss and central scotoma with reduced FMERGs were consistent with transient macular ischemia, and prophylactic anti-coagulation treatment was considered.

## Comment

Our case had an acute monocular visual loss and a fundus that appeared normal except for the relative delayed angiographic retinal filling time and enlarged disc cups. The RAPD, central scotoma, and the systemic complications made it difficult to determine the site of the alterations. The reduced FMERGs pointed to the retina as the site.

However, an abnormally long-lasting visual decrease is not typical for amaurosis fugax,[1] and ophthalmoscopy did not show retinal edema typical for arteriolar occlusion. Conventional electrophysiological examinations such as full-field ERGs and VEPs might be useful except when the ischemic site is in the macular region.

The clinical course in our case was compatible with a transient retinal ischemia with a possibility of a transient central retinal artery occlusion although additional influence of more generalized abnormality cannot completely be excluded. The ERG from 5 degrees area is similar to that of monkey ERG treated by 2-amino-4-phosphonobutyric acid (APB) and

cis-2,3-piperidine dicarboxylic acid (PDA) to suppress both on and off synapses[4,5]. This implies that the inner retinal layers have serious dysfunction. Because of the systemic complications, the risk of utilizing prophylactic anticoagulant agents was discussed. However, the VA and central scotoma quickly recovered accompanied by an improvement in the FMERGs without any intervention.

These findings indicate that clinicians should consider focal macular dysfunction in cases of acute vision loss and normal retinal appearance, and multifocal ERGs or FMERGs are useful in determining the site of the pathology.

## Acknowledgements

Support of this study was provided by Suzuken Memorial Foundation. No author has a financial or proprietary interest in any material or method mentioned.

## References

1. Joseph F. Rizzo III. Neuroophthalmologic Disease of the Retina. In Daniel MA (ed): *Principles and Practice of Ophthalmology, 2nd ed.* Philadelphia, W.B. SAUNDERS COMPANY, 2000: 4083-108.

2. Miyake Y, Shiroyama N, Horiguchi M, *et al.* Oscillatory potentials in electroretinograms of the human macular region. *Invest Ophthalmol Vis Sci* 1988;29: 1631-5

3. Miyake Y. Focal macular ERGs. In Miyake Y (ed): *Electrodiagnosis of Retinal Diseases.* Tokyo, Springer-Verlag, 2005: 20-32

4. Miyake Y. On and Off Responses in photopic ERGs. In Miyake Y (ed): *Electrodiagnosis of Retinal Diseases.* Tokyo, Springer-Verlag, 2005: 99-100

5. Sieving PA, Murayama K, Naarendorp F. Push-pull model of the primate photopic electroretinogram: a role for hyperpolarizing neurons in shaping the b-wave. *Vis Neurosci* 1994;11:519-32.

## Figure Legends

**Figure 1.** Fundus photograph and fluorescein fundus angiograms of the patient's left eye.

**Top left:** Fundus photograph showing enlarged cupping of the left optic disc.

**Top right:** Fluorescein angiogram with delayed arm-to-retina time.

**Middle left and right:** Fluorescein angiogram showing neither localized arteriolar filling delay nor retinal emboli in the middle (Left) and late (Right) phases.

**Bottom left:** Goldmann visual field in the acute phase shows small central scotoma and peripheral constriction in addition to a decrease in the isopters in the left temporal inferior area (quadrantic homonymous hemioanopsia) which had existed for years in the left eye. The central scotoma was new. Visual acuity was 20/2000 OS.

**Bottom middle:** Goldmann visual field performed six months earlier showing a decrease in the isopters in the left nasal inferior area (homonymous hemioanopsia ) in the right eye. Visual acuity was 20/30 OD.

**Bottom right:** Goldmann visual field performeymiyake@asu.aasa.ac.jpd three months after the onset showing disappearance of the central scotoma and an improvement of peripheral isopters. Visual acuity was 20/30 OD.

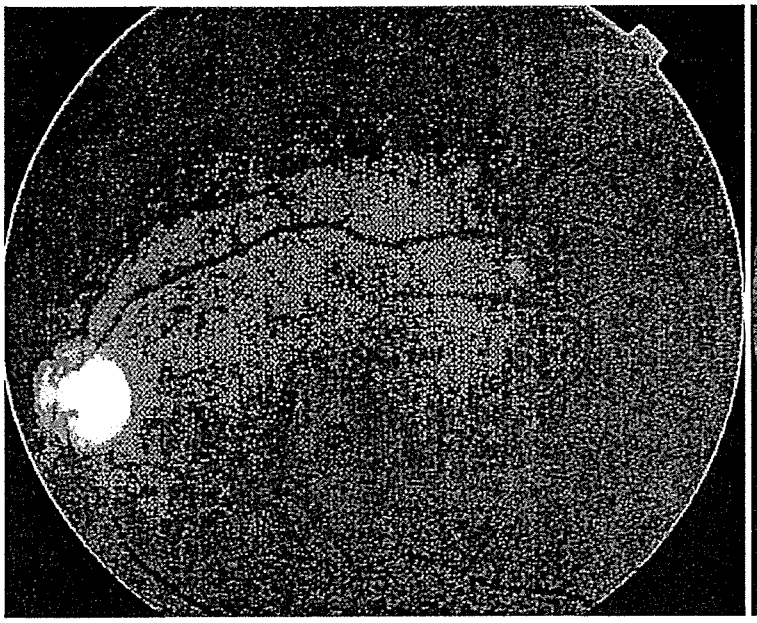
**Figure 2.** Focal macular electroretinogram (ERG, **Top**), full field electroretinogram and visual evoked response (VEP) from both eyes (**Bottom**).

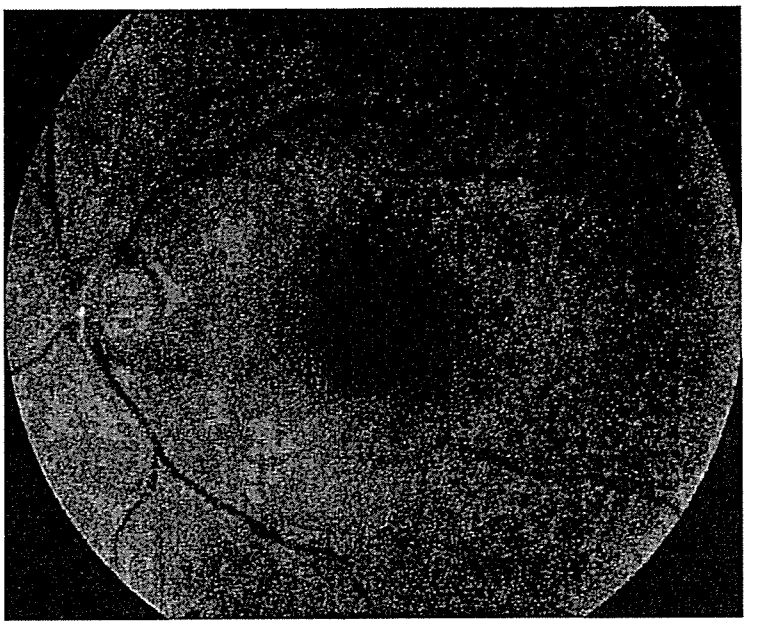
**Top:** Upper three recordings show the photopic a- and b-waves, and lowest recordings show the oscillatory potentials (Ops). The stimulus spot size was changed from 5 degree, 10 degree, and 15 degree as indicated in the figure. Left and Center: Focal macular ERGs recorded during the acute phase. The ERGs of the left eye (Center) are reduced compared with those of the of right eye (Middle). Focal macular ERGs recorded on the following day (Right) showing the recovery of all components of the left eye. The a-wave is not affected as much as the b-waves and Ops. **Bottom left**: Bright-flash ERG showed no significant difference in the two eyes even in the oscillatory potentials which are sensitive to the retinal ischemia. **Bottom right**: The VEPs showed no significant difference between the two eyes. The stimulus onset of full field ERG and VEP is indicated by arrow.

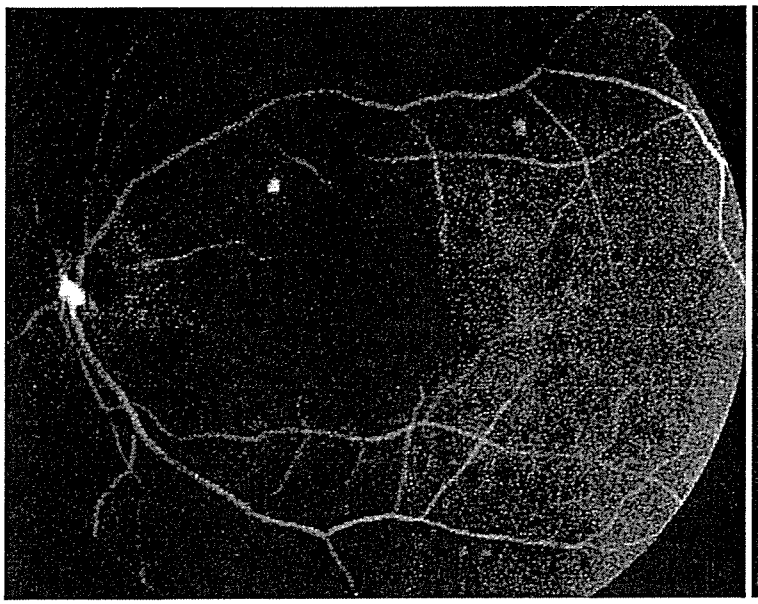
**Legend for Supplemental figure**

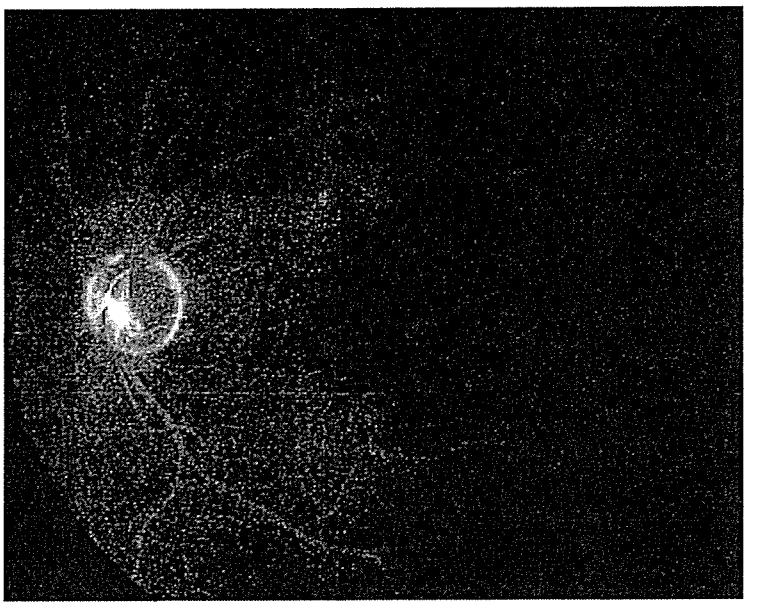
Left: The appearance of recording focal macular electroretinograms.

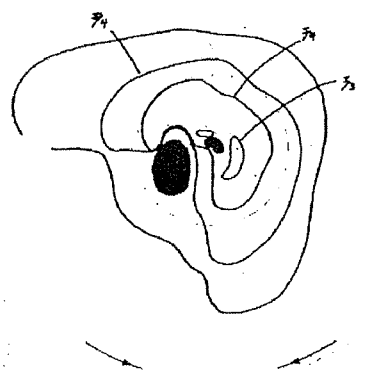
Right: Stimulating spot is superimposed on the fundus photograph. An infrared camera integrated into the fundus camera enables stimulating the macular spot under direct observation. Recording was performed under dim light condition. The luminance of the white stimulus spot and background light was 180.0 candelas [cd]/$m^2$ and 5000.0 cd/$m^2$, respectively. The FMERGs were elicited by a 5-Hz rectangular stimulus (equal on and off) centered on the fovea. The responses were differentially amplified and bandpass filtered (50 to 500 Hz for a- and b-waves; 5 to 500 Hz for oscillatory potentials). A total of 500 responses were averaged.

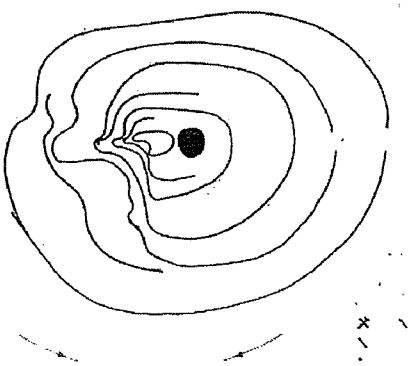


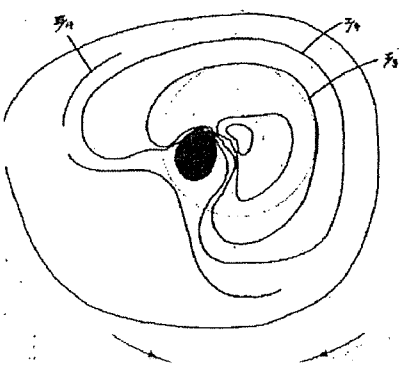












シンポジウムⅠ
「視機能を客観的に評価する」

# 網膜神経活動のイメージング
## ―網膜内因性信号計測法―

角　田　和　繁

東京医療センター臨床研究センター視覚生理学研究室

# A novel imaging technique of neural function in retina
## -Intrinsic signal imaging of retina-

Kazushige Tsunoda

Laboratory of Visual Physiology National Institute of Sensory Organs

要　約

視覚的機能を他覚的に評価することは、眼疾患の早期発見および治療効果の判定のために基本的かつ重要な課題である。そのために眼科における画像診断技術（イメージング）は近年めざましい進歩をとげてきたが、それらは主に解剖学的構造の把握を目的としており、視細胞をはじめとする網膜の神経活動を捉えることはできない。我々が開発した網膜内因性信号計測法は、非侵襲的な網膜神経機能のイメージング法であり、神経活動にともなって神経組織の微細構造や血流が変化する現象を利用したものである。この技術により、従来は不可能であった錐体視細胞、杆体視細胞の機能的マッピングを、高い解像度で行うことができるようになった。動物実験における信号の閾値は網膜電図（ERG）のb波とほぼ同等であることが示されており、現在ヒトでの計測が可能なように開発を行っている。このイメージング法が実現すれば、様々な網膜疾患において精度の高い他覚的機能評価が可能になると期待されている。

別冊請求先（〒152-8902）東京都目黒区東が丘２－５－１
東京医療センター　臨床研究センター視覚生理学教室　角田　和繁
Key words：網膜内因性信号計測法、イメージング、網膜機能評価

## ＜はじめに＞

眼科における画像診断技術は近年めざましい進歩をとげてきた。たとえば光干渉断層計（OCT）は、検眼鏡によって捉えることのできない網膜微細構造の観察を可能にするものであり、網膜疾患の診断、治療に関する従来の常識を一変させるほど臨床応用価値の高いものである。しかしOCTや、走査型レーザー検眼鏡（SLO）などの画像診断法は、解剖学的構造の把握を目的としており、これによって視細胞をはじめとする網膜の神経活動を捉えることはできない。したがって、網膜機能（神経活動）の他覚的評価のためには、電気生理学的検査である網膜電図（ERG）が今でも主要な役割を果たしている。

網膜、なかでもその中心に位置する黄斑部は、視力の維持のために重要な部位であり、同時に臨床的に様々な疾患が起こりやすい部位でもある。我々のグループでは、ERGとはまったく異なるしくみで網膜の神経活動を非侵襲的にイメージングする方法を開発し、新たな眼科検査法としての実用化に向けた研究を行っている（網膜内因性信号計測法、Functional Retinography）。内因性信号計測法が眼底における機能計測に応用できれば、網膜神経機能の客観的な評価法として疾患の早期発見や早期治療につながることが期待される。本項では網膜内因性信号計測法がいかなる検査法であるか、またこの検査法の展望について紹介したい。

## ＜網膜内因性信号計測法とは＞

神経活動にともなって神経組織の微細構造や光反射率が変化する現象は古くから知られている[1]。特に、神経活動にともなう代謝変化を光の反射率変化として捉える計測法は内因性信号計測法（Intrinsic Signal Imaging）と呼ばれ、最近の脳科学の進歩に大きく貢献してきた[2-5]。実際の計測法は、神経組織をCCDカメラでイメージングし、刺激前と刺激後の画像を重ね合わせて比較するという非常に単純なものである。刺激後に画像の明るさが変化している部分が神経活動の起きた領域に相当し、通常は神経活動の高い領域がより暗く見える。信号の起源として、神経活動にともなう光散乱変化や血中ヘモグロビン飽和度の変化などが考えられている（図１）。

しかし、大脳皮質において光学計測法を行うためには脳表面を露出する必要があり、比較的大きな侵襲を伴うものである。したがってその臨床医学への応用は、てんかん患者に対する脳外科手術時に行われる実験的なものに過ぎない[6]。一方、ヒトの網膜は内因性信号の計測に有利な特徴をいくつか備えている。まず、眼球の光学系（角膜、水晶体）を通して網膜を直接観察することができ、眼球自体が光学計測における理想的な観察用チェンバーとして働くため、手術的な侵襲が必要ない。さらに、網膜外層では視細胞が密に一定方向に並んでいるため、光散乱変化による反射率変化をとらえやすい。筆

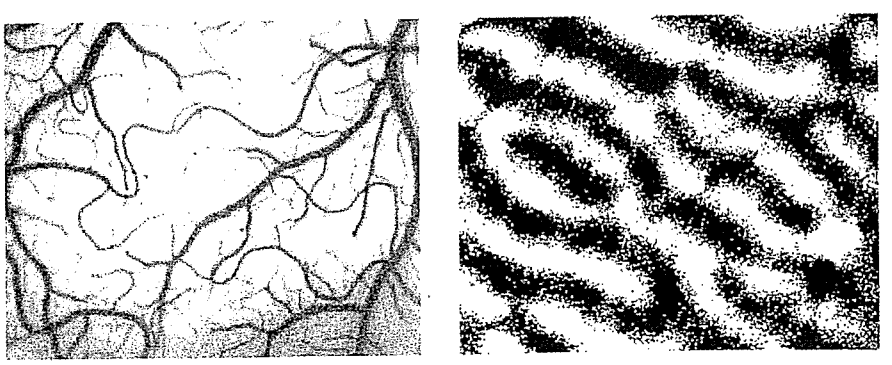


図１（左）アカゲサル第一次視覚野の皮質表面。
（右）内因性信号計測によって描出された同部位の眼優位性コラム。黒い帯が右眼の入力を受けて神経活動が高まっている領域。

者はこの技術の眼科分野への応用に着目し、早くから網膜における実験を行ってきた[7]。

## ＜測定方法＞

開発のための動物実験では、ヒトとほぼ同じ解剖学的構造を持つニホンザル、アカゲザルの眼底を用いている。

全身麻酔下において非動化した眼底を、眼底カメラを改良した観察系を用いてCCDカメラでモニターする（図２）。眼底観察用のハロゲン光は、干渉フィルター（800－900nm）を透過して眼底後極部を照明する。解像度640×480、毎秒30フレームのCCDカメラによって眼底からの光反射率変化を持続的に記録する。測定開始から0.5秒後に眼底後極部全体を白色キセノンフラッシュ（１ms）にて刺激する。一回の測定は通常５－10秒間行う。

刺激前0.5秒間の平均画像の反射率と、刺激後の画像における反射率との比をピクセル毎に計算し、その比を256階調にスケーリングし画像化する。大脳皮質での測定では通常10－30回ほどの加算を要するが、網膜においては加算を行わない、single trialでのマッピングが可能である。

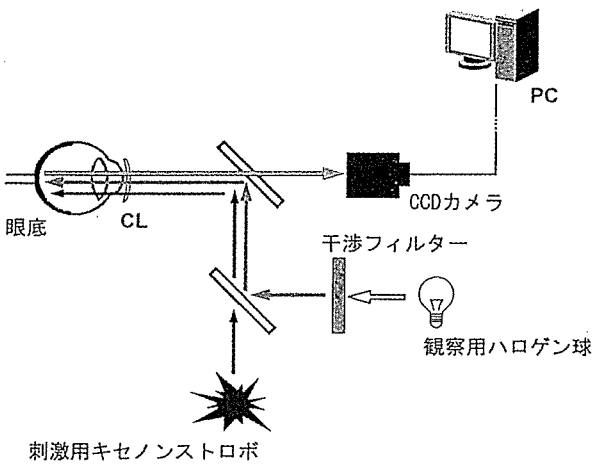



図２　実験セットアップの模式図。

## ＜結　　果＞

フラッシュによるびまん性刺激によって視細胞が活動すると、網膜全体の反射率が低下し画像では暗く描出される（図３）。この内因性信号は刺激後150msにピークを持つ早い反応で、中心窩で最も強い。信号強度を疑似カラーで表示すると、明順応下では中心窩に内因性信号の急峻なピークを認め、周辺部に向かって減少す

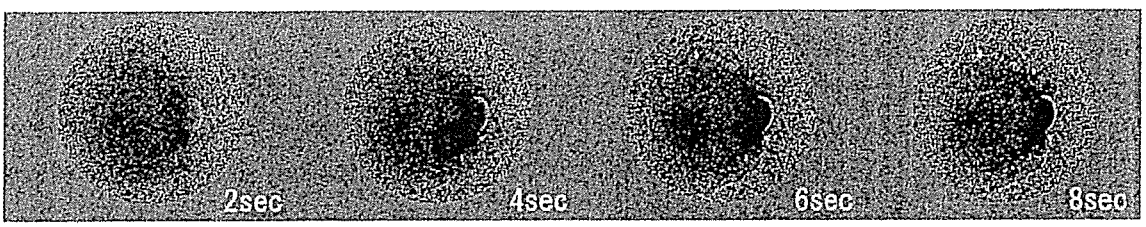



図３　赤外光観察下（800－900nm）における網膜内因性信号。後極部が暗く変化しているのが分かる。

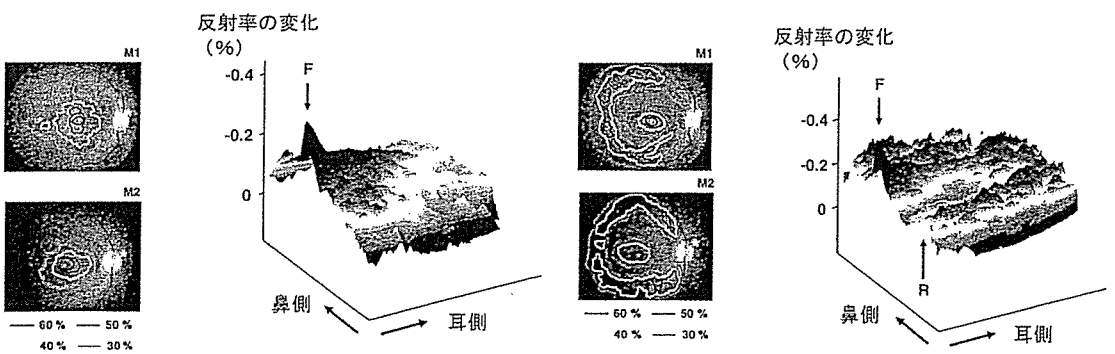



図４　網膜内因性信号のトポグラフィー。左図は明順応下、右図は暗順応下における記録。等高線は反応のピークに対する相対値を示す。
（白点：中心窩、＊印：視神経乳頭部、Ｆ：Fovea、Ｒ：Rod ring）

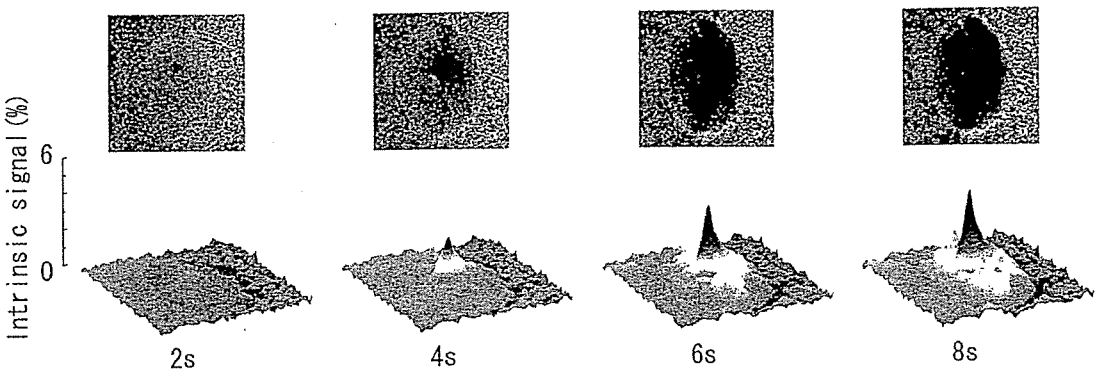



図5　視神経乳頭部におけるフラッシュ刺激後の内因性信号変化（上）および同部位のトポグラフィー（下）。

るが、暗順応下では中心窩に加えて周辺部にドーナツ状のピークを認める（図4）。内因性信号のピークは中心窩では錐体視細胞に、周辺部では杆体視細胞の解剖学的な分布（rod ring）によく一致しており、網膜内因性信号の発生には視細胞が大きく寄与していると思われる[8-9]。

さらに、視神経乳頭部に注目すると、ここでは網膜面とは異なりフラッシュ刺激後にゆっくりと信号が強くなっていく。これは刺激後の血流増加を反映した光散乱強度変化と考えられ[10]、中心動静脈に相当する視神経乳頭中央部に急峻なピークが見られる（図5）。

なお、同一の刺激系を用いて角膜電極より暗順応ERGを記録し、内因性信号と比較すると、黄斑部を除く網膜面および視神経乳頭における内因性信号の閾値は、ERGのb波とほぼ同一であることが分かった。これは、網膜内因性信号計測法が非常に感度の高い神経機能計測法であることを示している。

## ＜ま と め＞

内因性信号計測法の利点は、赤外光の反射率変化を計測するために非侵襲的であること、空間分解能が高いこと（理論上は眼底写真に相当）、測定時間が短いことなどである。問題点としては、ヒトの測定時に生じる固視微動等により、画質が著しく低下することであり、現在それを克服すべく研究を行っている。将来臨床応用が可能になれば黄斑変性症や網膜色素変性症など様々な網膜疾患において精度の高い他覚的機能評価が可能になると期待されている。

さらに網膜内因性信号計測法以外にも、近年はOCTを利用して神経機能評価を行おうという研究も行われている。これは2002年に理化学研究所のMaheswariらによって初めて提唱されたFunctional OCT[11]という概念であり、網膜における応用に向けた研究がすでにいくつかの国の研究機関で行われている。網膜の神経機能をイメージングするという研究は新しい診断法として高く注目されており、将来はERGと並ぶ新たな客観的機能評価法が確立される日が来るに違いない。

### 参考文献

1）Cohen L. Changes in neuron structure during action potential propagation and synaptic transmission. Physiol Rev. 53: 373-418, 1973

2）Grinvald A, Lieke E, Frostig RD, Gilbert CD, Wiesel TN. Functional architecture of cortex revealed by optical imaging of intrinsic signals. Nature, 1986; 324: 361-364.

3）Bonhoeffer T, Grinvald A. Optical Imaging Based on Intrinsic Singals: The Methodology. In: Toga AW, Mazziotta JC, eds. Brain Mapping. San Diego: Academin Press; 1996: 55-97.

4）Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. Nat Neurosci.

2001; 4: 832-8.

5) Fukuda M, Maheswari RU, Homma R, Matsumoto M, Nishizaki M, Tanifuji M. Localization of activity-dependent changes in blood volume to submillimeter-scale functional domains in cat visual cortex. Cereb Cortex. 2005 Jun; 15 (6): 823-33.

6) Haglund MM, Ojemann GA, Hochman DW. Optical imaging of epileptiform and functional activity in human cerebral cortex. Nature 1992 358: 668-671.

7) Tsunoda K, Oguchi Y, Hanazono G, and Tanifuji M. Mapping Cone- and Rod-Induced Retinal Responsiveness in Macaque Retina by Optical Imaging. Invest. Ophthalmol. Vis. Sci. 2004 45: 3820-3826.

8) Osterberg G. Topography of the layer of rods and cones in the human retina. Acta ophthalmol. 1935; 13: 6-97.

9) Curcio CA, Sloan KR, Jr., Packer O, Hendrickson AE, Kalina RE. Distribution of cones in human and monkey retina: individual variability and radial asymmetry. Science. 1987; 236: 579-82.

10) Riva CE, Falsini B, Logean E. Flicker-evoked responses of human optic nerve head blood flow: luminance versus chromatic modulation. Invest Ophthalmol Vis Sci. 2001; 42: 756-62.

11) Maheswari RU et al. Implementation of optical coherence tomography (OCT) in visualization of functional structures of cat visual cortex. Opt. Comm. 202: 47-54, 2002

*J Neurophysiol* 96: 3147–3156, 2006.
First published August 30, 2006; doi:10.1152/jn.01224.2005.

# Representation of the Spatial Relationship Among Object Parts by Neurons in Macaque Inferotemporal Cortex

**Yukako Yamane,[1,2] Kazushige Tsunoda,[1,3] Madoka Matsumoto,[1] Adam N. Phillips,[1] and Manabu Tanifuji[1]**
[1]*Laboratory for Integrative Neural Systems, RIKEN Brain Science Institute, Saitama;* [2]*Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo;* [3]*Laboratory of Visual Physiology, National Institute of Sensory Organs, Tokyo, Japan*

Submitted 21 November 2005; accepted in final form 27 August 2006

**Yamane, Yukako, Kazushige Tsunoda, Madoka Matsumoto, Adam N. Phillips, and Manabu Tanifuji.** Representation of the spatial relationship among object parts by neurons in macaque inferotemporal cortex. *J Neurophysiol* 96: 3147–3156, 2006. First published August 30, 2006; doi:10.1152/jn.01224.2005. We investigated object representation in area TE, the anterior part of monkey inferotemporal (IT) cortex, with a combination of optical and extracellular recordings in anesthetized monkeys. We found neurons that respond to visual stimuli composed of naturally distinguishable parts. These neurons were sensitive to a particular spatial arrangement of parts but less sensitive to differences in local features within individual parts. Thus these neurons were activated when arbitrary local features were arranged in a particular spatial configuration, suggesting that they may be responsible for representing the spatial configuration of object images. Previously it has been reported that many neurons in area TE respond to visual features less complex than natural objects, but it has remained unclear whether these features are related to local features of object images or to more global features. These results indicate that TE neurons represent not only local features but also global features such as the spatial relationship among object parts.

## INTRODUCTION

Visual information about object images is conveyed from early visual area V1 to inferior temporal (IT) cortex through areas V2 and V4 in macaque monkeys (for review, see Logothetis and Sheinberg 1996). Area TE is a ventral part of IT cortex and is the site that represents object images necessary for visual recognition (Gross 1994; Logothetis and Sheinberg 1996).

Early studies on visual responses of TE neurons showed that these neurons respond to various visual stimuli including natural object images (Bruce et al. 1981; Desimone et al. 1984; Gross et al. 1979; Perrett et al. 1982; Schwartz et al. 1983). More recently, a number of studies have attempted to identify the simplest visual features that activate individual neurons in area TE (Kobatake and Tanaka 1994; Tanaka et al. 1991). These studies have revealed that essential stimuli for TE neurons are visual features that are geometrically less complex than natural objects. Thus combinations of visual features are necessary for neural representation unique to individual object images in area TE.

As in the primary visual cortex, neurons in area TE with similar response properties are reported to be clustered into columns (Fujita et al. 1992; Gochin et al. 1991). The columns responding to visual stimuli have been visualized with intrinsic signal imaging as darkened spots scattered across the cortical surface (Tsunoda et al. 2001; Wang et al. 1996, 1998). In particular, Tsunoda and colleagues used this technique together with conventional extracellular recordings and showed that an object image activates multiple spots, each of which represents a particular visual feature of the object image (Tsunoda et al. 2001). They reported that some of the visual features represented by activated spots were local features of object images, that is, features that appear in a spatially localized part of the object image. Thus it remains unknown how spatial arrangements of these local features in an object image are specified. Neurons in area TE may represent global features, such as spatial configuration of local features, in addition to spatially localized features. In this paper, we address this question by combining data from intrinsic signal imaging and extracellular recordings.

## METHODS

### *Anesthesia and the general recording condition*

Four rhesus monkeys were artificially ventilated with a mixture of $N_2O$, $O_2$, and isoflurane for anesthesia and paralyzed with pancuronium bromide or vecuronium bromide (Tsunoda et al. 2001). The visual stimuli were presented on a 20-in CRT display placed 57 cm from the eye contralateral to the recording hemisphere. The pupil of the eye was dilated by local application of 0.5% tropicamide 0.5% phenylephrine, and the cornea was covered with a contact lens of appropriate power to focus the visual stimuli onto the retina. The fovea was identified with a custom-made ophthalmoscope, and the position of the fovea was back-projected onto the center of the CRT screen. Except for three-dimensional (3D) objects for manual presentations, the visual stimuli were presented at the center of the CRT display. Electroencephalography (EEG), electrocardiography (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 Institute of Health.

### *Intrinsic signal imaging*

The dorsal part of area TE was exposed and illuminated by light with a wavelength of 605 nm through a glass cover slip window attached to a titanium chamber centered 15.0–17.5 mm anterior to the ear bar position (Tsunoda et al. 2001). Reflected light from the cortex was detected by a low-noise video camera (frame rate, 1/30 frames/s; S/N ratio, 60 dB; CS8310, Teli, Japan) and digitized by a 10-bit video capture board (Pulsar, Matrox). The light was focused to a depth of 500 $\mu$m below the cortical surface. The imaged area was 6.5 × 4.9

Address for reprint requests and other correspondence: M. Tanifuji, Laboratory for Integrative Neural Systems, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan (E-mail: tanifuji@riken.jp).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

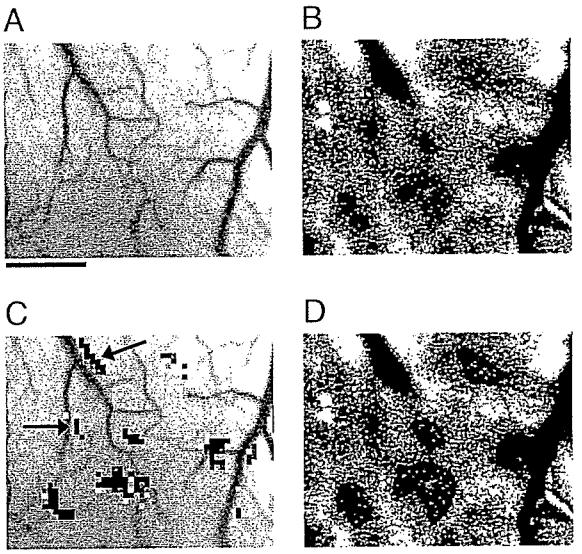



FIG. 1. Intrinsic signal imaging detected local modulation of light absorption changes in area TE. *A*: surface view of the exposed portion of the cortex. *B*: differential image showing local increase in light absorption. *C*: regions that showed statistically significant darkening after presentation of a visual stimulus are indicated in red ($P < 0.05$). Apparent artifacts that appeared along the thick vessels (arrows) were eliminated from the analysis. *D*: extracted active spots outlined by connecting pixels with half the peak absorption value. Scale bar = 1 mm.

mm and contained 320 × 240 pixels. We presented a visual stimulus to the monkey for 2.0 s, and sequential images were acquired for 4.0 s (starting from 1.0 s before the stimulus onset). During the 2-s stimulus presentation period, a stimulus image appeared and moved in a circular path (with a radius of 0.4° at the rate of 1 cycle/s). The imaging experiments consisted of two sessions. In the first session, the visual stimuli were 10–20 object images together with two blank images as control. Then on the basis on these results, we selected several stimuli that activated a large number of spots in the imaged region. In the second session, the selected stimuli ("the original"), their modifications and two controls were used as visual stimuli. Each stimulus was randomly presented 15–30 times in one session. The same imaging session as the second session was repeated at least twice on different days to confirm the consistency of the observed spots.

### *Identification of the active spots*

The active spots were extracted as follows (Tsunoda et al. 2001): *1*) images acquired during the 0.5- to 3.0-s period after the onset of stimulus presentation were divided by an average of images during the 1-s period just before the stimulus onset. *2*) Gaussian spatial filtering was used to eliminate the global (stimulus nonspecific) darkening and high-frequency noise (cut-off frequencies: $\sigma = 0.04$ mm$^{-1}$ for high cut and $\sigma = 2.1$ mm$^{-1}$ for low cut). *3*) The *t*-values were calculated by pixel-by-pixel comparison of signal intensity between the filtered images for the trials with a particular stimulus and those for the control trials. The filtered images with a stimulus were averaged for all the trials and a differential image was created by subtracting the averaged image for control trials. Localized dark regions of the differential image, which showed significant darkening (*t*-test, $P < 0.05$), were defined as active spots. *4*) The contour of active spots was demarcated at the half-value of the peak absorption value. Representative images for each step are shown in Fig. 1.

### *Extracellular recording*

The exposed cortex used for intrinsic signal imaging was covered with a transparent artificial dura made of silicon rubber (Arieli et al. 2002). Tungsten microelectrodes were inserted into the spots through the artificial dura. The surface blood vessel pattern was used as a mapping reference to identify the position of the spots. Extracellular action potentials were recorded for 3 s in each trial. Visual stimulus presentation started 1 s after the onset of a trial and lasted for 1 s. During the 1-s stimulus-presentation period, a stimulus image appeared and moved in a circular path (with a radius of 0.4° at the rate of 1 cycle/s). No intertrial interval was inserted, so that a blank period between two stimuli was 2 s. The different stimuli were presented in pseudo-random order, and the number of trials for each stimulus was between 10 and 20. For each stimulus, we applied the Wilcoxon test to the difference in the mean firing rate during and before the stimulus presentation. The amplitude of evoked responses for each stimulus was calculated by subtracting the mean firing rate during the 1-s period before the stimulus onset from the mean firing rate during the 1-s stimulus-presentation period, and by averaging for all the trials.

To characterize individual cells, we determined visual features critical for the cells according to previous studies (Fujita 1993; Fujita et al. 1992; Tanaka et al. 1991) (Fig. 2): *1*) we manually searched for the most effective visual stimulus among 96 hand-held 3D objects (Fig. 3), *2*) we simplified the best stimulus by removing or modifying a particular visual feature of the stimulus, and *3*) if the simplified image elicited significant responses (Wilcoxon test, $P < 0.05$) and also if the response amplitude for the simplified image exceeded a certain threshold, we used this image as the best stimulus in the next step. This procedure was repeated until further simplification failed to produce any response that exceeded the threshold. The threshold was set to 70% of the response elicited by the stimulus before simplification because there was no significant difference in evoked responses at this threshold. Typically, we started with the examination of a

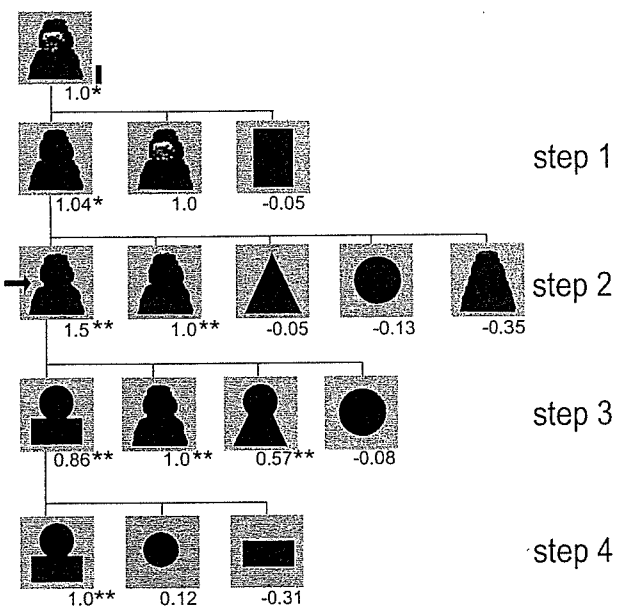



FIG. 2. Systematic simplification of an object image. The stimulus was simplified step by step. The stimulus that evoked the strongest response in one step was examined in the next step as the reference stimulus. The numbers below each picture indicate the response amplitudes normalized to the response to the reference stimulus together with statistical significance (Wilcoxon test, $*P < 0.05$, $**P < 0.01$). Step 1 showed that neuronal activities elicited by the best object and the silhouette were the same. Step 2 examined the effect of the 'sharpness' of the corner at the junction of upper and lower parts (arrow) and showed that the silhouette with the sharpest corners (leftmost picture) was the most effective stimulus. Step 3 showed that activities elicited by the leftmost two stimuli were not significantly different. Step 4 showed that neither the upper nor lower part activated the cell. In this case, the critical feature was determined as a combination of a circle and a rectangle (leftmost picture at step 4). Scale bar, 5°.

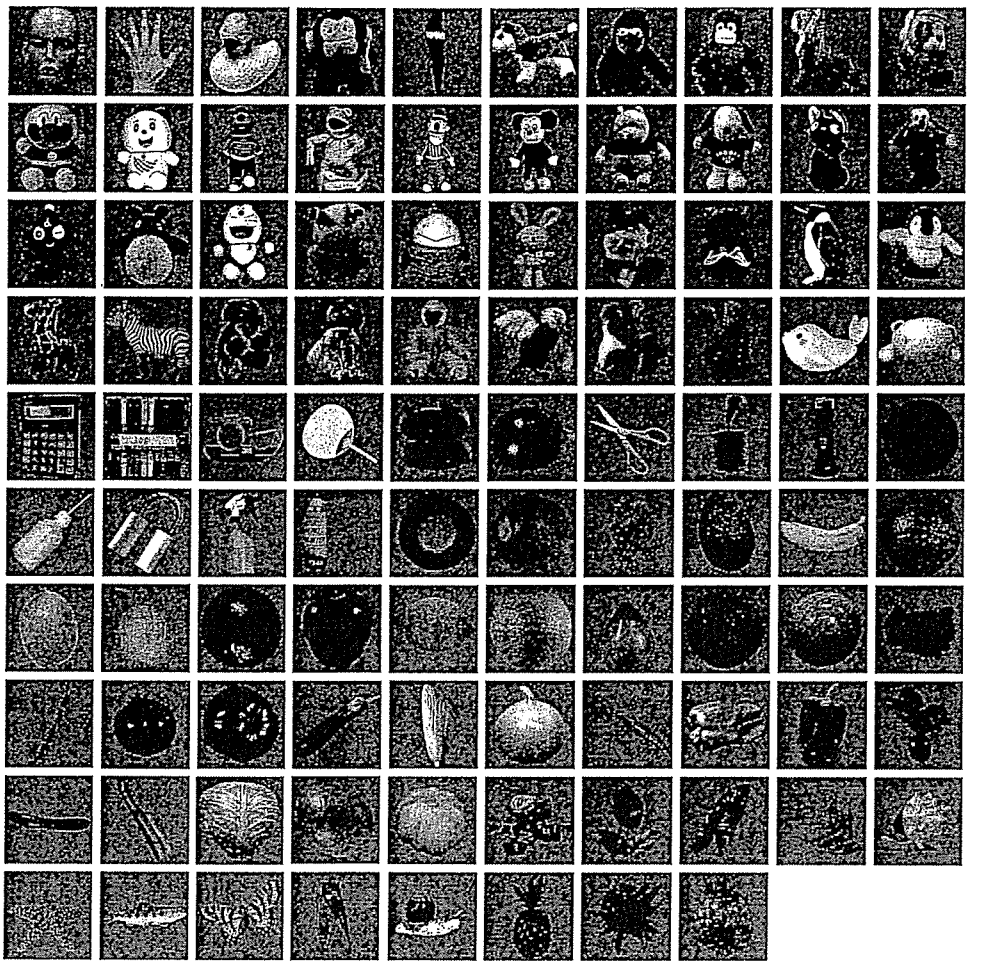


FIG. 3. The 3-dimensional (3D) objects used as visual stimuli to search for effective objects for individual cells. These objects were presented to the animals from various perspectives, as shown in the last three pineapple images in the *bottom row*.

monochrome image and silhouette of the original as in Fig. 2. However, image simplifications in the intermediate levels were different from case to case even if the original object was the same. The average numbers of simplification steps before reaching the simplest visual feature was 4.8 ± 2.0 (mean ± SD).

We examined receptive field size of each cell manually with the most effective 3D objects. We presented these objects not through the CRT display but directly to the animals because the size of receptive fields was larger than the size of the CRT display in most of the cells. On average, edges of receptive fields were 18.82 ± 8.17° above, 23.14 ± 5.30° below, 23.86 ± 8.41° contralateral, and 22.60 ± 8.62° ipsilateral to the recording hemisphere (means ± SD, $n = 35$). These values were measured from the center of the fovea. These values represent distances between the center of the CRT screen and the center of objects at the edges of receptive fields. Two exceptional cell having smaller receptive field size was eliminated from the analysis. The size of the object images used for quantitative analyses was on average 13.30 ± 3.08° along the vertical axis and was 10.88 ± 4.02° along the horizontal axis.

### *Definition of object parts*

Throughout this manuscript, we took the simplest definition of object parts as the ones naturally distinguishable by discontinuities at minima of negative curvature of the object shape. For example, the minimum of negative curvature of stimulus 1 in Fig. 4*A* is the joint connecting head and body, and accordingly the stimulus is segmented into "head" (stimulus 3) and "body" (stimulus 2). Although this definition of parts is conventionally used in the field of object vision, there is no a priori reason to define object parts according to this definition for TE neurons. Our intention, however, was not to explore the optimal segmentation for TE neurons, but rather to search for a possible mechanism for representing the spatial relationship among local features.

## RESULTS

### *Intrinsic signal imaging*

First, using intrinsic signal imaging, we identified one or two visual stimuli that activated a large number of spots in an imaged region of area TE. Each of these "original" visual stimuli was segmented into parts containing local features. We then conducted another intrinsic signal imaging session with a stimulus set consisting of the original (stimulus 1), two individual parts of the original (stimuli 2 and 3), and the original with a gap between the two parts (stimulus 4; Fig. 4). The results revealed that each spot was activated differently by these four stimuli. For example, one spot (indicated by an arrowhead in Fig. 4*A*) was activated by stimuli 1, 3, and 4 but not by stimulus 2. We interpret activity in this spot as being related to local features in the upper part of the original image (Fig. 4*A*). Because we were interested in identifying spots related to the spatial arrangement of local features in parts, we restricted our analysis to spots that were activated by combinations of the two parts but not by individual parts. Among four examined hemispheres, we found three spots, A–C, that satisfied this criterion: these spots were activated by the original image (stimulus 1) and the original with a gap (stimulus 4)

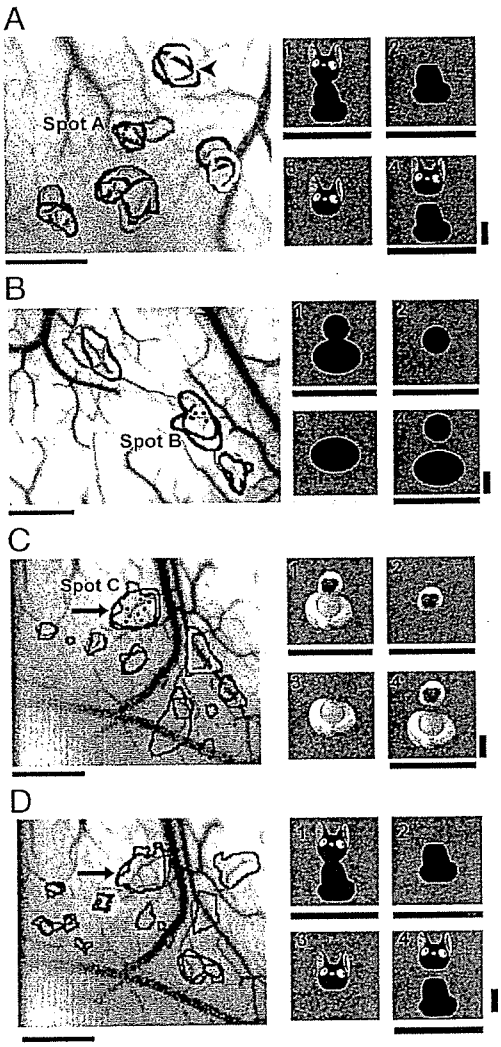



FIG. 4. Activation patterns of intrinsic signals evoked by sets of visual stimuli. *A–C*: activation patterns obtained from 3 different hemispheres, in which spots A, B, and C [spatial relationship relevant spots (SRR spots)] had specific response selectivity to sets of visual stimuli. *D*: activation pattern obtained from the same area of cortex shown in *C* but with a different set of stimuli. In each panel, activation patterns evoked by a set of the stimuli 1–4 (right) are indicated by colored contours superimposed on the surface image of the cortex. The bars below each stimulus and the outlines of spots activated by the stimulus match in color. The arrowhead in *A* indicates the spot related to local features in stimulus 3. The arrow in *D* indicates the region corresponding to spot C. The dots in the spots A–C indicate electrode penetration sites for subsequent extracellular recordings. Horizontal scale bar, 1 mm. Vertical scale bar, 5°.

but not by either part alone (stimuli 2 and 3; Fig. 4). Because the original with a gap (stimulus 4) activated these spots, activity in these spots could not be caused by specific responses to particular local features at the junction between the parts, such as a sharp negative curvature. We therefore considered these spots to be spatial relationship relevant spots (SRR spots), where we would likely find neurons representing the spatial relationship between two parts or between features within these parts.

### *Responses of the cells in SRR spots to spatial arrangements of object parts*

We then conducted extracellular recordings from 49 cells located within SRR spots to characterize responsiveness of individual cells (13, 14, and 22 cells in spots A–C, respectively). First, we examined visual responses of each cell with 96 real object stimuli including faces, hands, imitations of living animals, stuffed animals, tools, and plastic fruits and vegetables (Fig. 3). These objects were presented in various sizes, orientations, and views so that the actual number of two-dimensional images used as visual stimuli was three or four times larger than the number of real objects. The stimuli that elicited significant responses (Wilcoxon test, $P < 0.05$) were diverse in color, texture, and local shapes (Fig. 5, *A, C,* and *E*). We could not explain this visual diversity in effective stimuli by preferred stimuli being different from cell to cell in a spot because individual cells in these spots responded to different stimuli and the response amplitudes did not significantly differ from each other (1-way ANOVA, $P > 0.25$; Fig. 5, *B, D,* and *F*). One common aspect of these effective visual stimuli was that the objects tend to consist of at least two distinguishable parts (Fig. 5, *B, D,* and *F*). These results from extracellular recordings were in accordance with the observation that with optical imaging, stimulus selectivity of a spot was the same for stimulus sets that originated from different object images (Fig. 4, *C* and *D*).

To address the question of whether cells in these spots could represent spatial relationships between object parts, we generated a set of visual stimuli for each cell where the upper part of the best object stimulus was rotated by various angles relative to the lower part of it. We then examined selectivity of each cell to this set of visual stimuli (Fig. 6). We found that cells were selectively activated when the parts were aligned vertically (Fig. 6, *A* and *B*). The selectivity could not be simply explained by changes in retinotopic positions of the upper parts that occurred incidentally during the spatial rearrangements of parts because of the large receptive field sizes of cells in area TE (Gross et al. 1969; Ito et al. 1995). In fact, evoked responses to the best object stimulus at different positions in space did not significantly differ from each other (Fig. 6*A*, symbols on the vertical axis). Of 30 cells examined for spatial arrangements of object parts, the responses of 20 cells (67%) significantly depended on the spatial configurations of two parts (1-way ANOVA, $P < 0.05$; Fig. 6, *C–F*). Thirteen cells had a single peak at 0 or 45° (Fig. 6, *C* and *D*), 4 cells had a single peak at other positions (Fig. 6*E*), and the remaining 3 cells had two peaks (Fig. 6*F*). Receptive field sizes of these cells were larger than the range of the shift of upper parts in retinotopic position incidental to the spatial rearrangement of parts (see METHODS). These results suggest that the cells in SRR spots are sensitive to particular spatial arrangements of objects' parts.

### *Response properties of the cells in SRR spots for the simplest visual features that activated the cells*

The sensitivity of the cells to a particular spatial arrangement of parts could be due to the changes in local shapes of either part that occurred incidentally during spatial rearrangements of parts (for example, see Fig. 6, *A* and *B*). We conjectured that this was not the case because SRR spots were less sensitive to variations in local shapes (Fig. 5). To confirm this point, however, we determined the simplest visual feature that produced maximal activation ("critical feature"), and examined the cell's sensitivity to modifications of the critical feature for

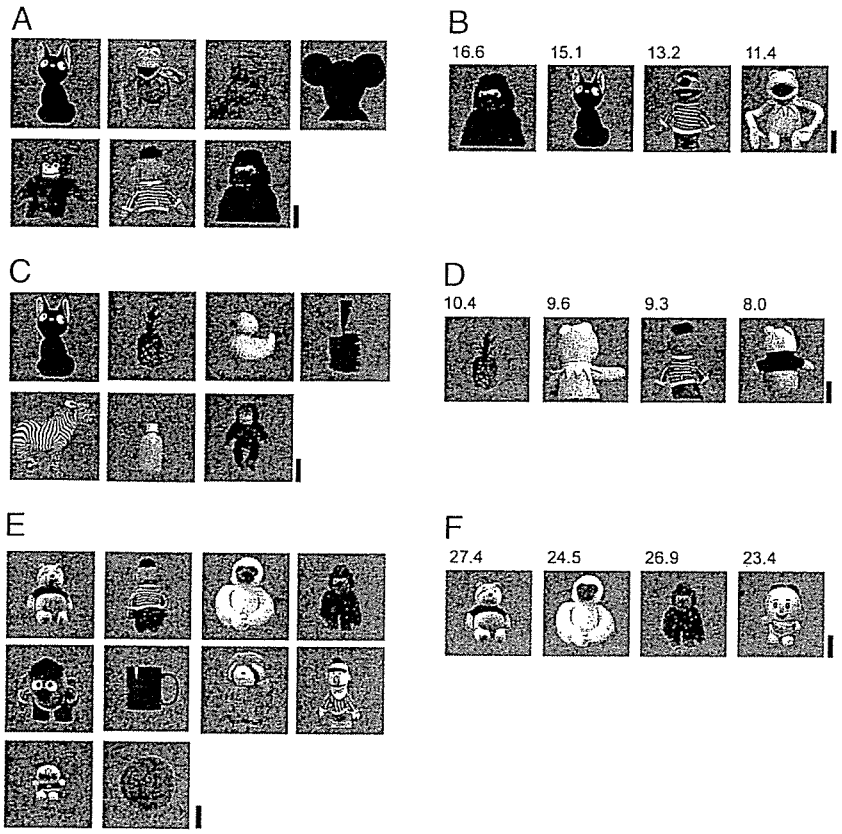



FIG. 5. Representative object stimuli that elicited significant responses of cells in SRR spots. *A*, *C*, and *E*: each object image represents the stimulus in a specific orientation that elicited the strongest significant responses out of the 96 objects in various orientations and views tested (*t*-test, $P < 0.05$) for a cell recorded in spots A (*A*), B (*C*), and C (*E*). *B*, *D*, and *F*: best 4 stimuli of 96 objects that elicited significant responses for a representative cell in spots A (*B*), B (*D*), and C (*F*). One-way ANOVA confirms no significant difference in responses to these 4 stimuli in both cases ($P > 0.25$). Evoked responses (spikes/s) are indicated above each stimulus image. Scale bar, 5°.

each cell in spots A and B. We systematically simplified the best object stimulus step by step to find critical features for 27 cells, following procedures from previous studies (Fujita 1993; Fujita et al. 1992; Kobatake and Tanaka 1994; Tanaka et al. 1991) (Fig. 2). Figure 7*A* shows the responses of a representative cell in spot A to its critical feature and to modifications of the critical feature. The critical feature was a combination of a circle and a rectangle (Fig. 7*A*, stimulus 1); the presentation of the upper or lower part alone caused significant decrease in the evoked responses (*t*-test, $P < 0.05$; Fig. 7*A*, stimuli 2 and 3). The cell responded equally well to the original colored object image and a silhouette of the original, indicating that color and texture of the stimulus were not essential (for example, see Fig. 2). The cell was not sensitive to changes in the shape of individual parts as long as the combination was preserved (Fig. 7*A*, stimuli 4 and 5). The existence of two parts was required, but the cell was not sensitive to local features at the junction between the two parts. For example, evoked responses to the stimulus with a gap (stimulus 7) and to the original (stimulus 1) were not significantly different (Fig. 7*A*). If, however, the "two parts" distinction was made less evident by smoothing the sharp joints (Fig. 7*A*, stimulus 8), the response was significantly reduced. Thus stimulus 6 but not stimulus 5 caused significant decrease in the evoked responses (Fig. 7*A*). The representative cell in spot B shows results consistent with the cell in spot A (Fig. 7*C*). In addition, we found that sharp joints between the two parts presented in isolation (Fig. 7*C*, stimulus 10) significantly reduced the responses, indicating that a sharp joint by itself was not sufficient.

Of 27 examined cells, the stimulus simplification revealed that 25 cells (93%) were not selective for color, luminance, or texture (exceptions are given in Fig. 7*B*, stimuli 5 and 6), and the critical features of 22 cells (81%) consisted of two parts (Fig. 7, *B* and *D*; exceptions are stimulus 8 in Fig. 7*B*, and stimuli 7 and 8 in Fig. 7*D*). The sensitivity of these neurons to modifications of critical features are summarized in the scatter plots, where diagonal lines indicate that evoked responses to the critical features and those to the modifications were the same (Fig. 8). Evoked responses to isolated parts were typically smaller than those to the critical features except for in a few cells (Fig. 8, *A* and *B*). The cells did not respond differently after changes in shape of either part or by inserting a gap between the two parts (Fig. 8, *C* and *D*). In 37% of the cells (filled symbols; 6 of 16 cells), significant reduction was observed when the stimuli had smoothed edges between the parts, but there was no significant reduction for the rest of the cells (open symbols; 10 of 16 cells) probably because two parts of the critical features were still distinguishable even when the joint was smoothed (Fig. 8*E*). Elongated shapes and a sharp joint between the two parts presented in isolation caused significant reductions in responses (Fig. 8, *F* and *G*). Thus sensitivity to these modifications of the critical features obtained from a population of neurons in these spots generally agreed with that obtained from representative cells.

In addition to the preceding results, as expected from the sensitivity of these cells to spatial arrangement of the object parts (Fig. 6), we found these cells were also sensitive to modifications of critical features in spatial arrangement of parts (Fig. 9). It should be noted that among the cells that were selective to a particular spatial arrangement of parts, 73% did

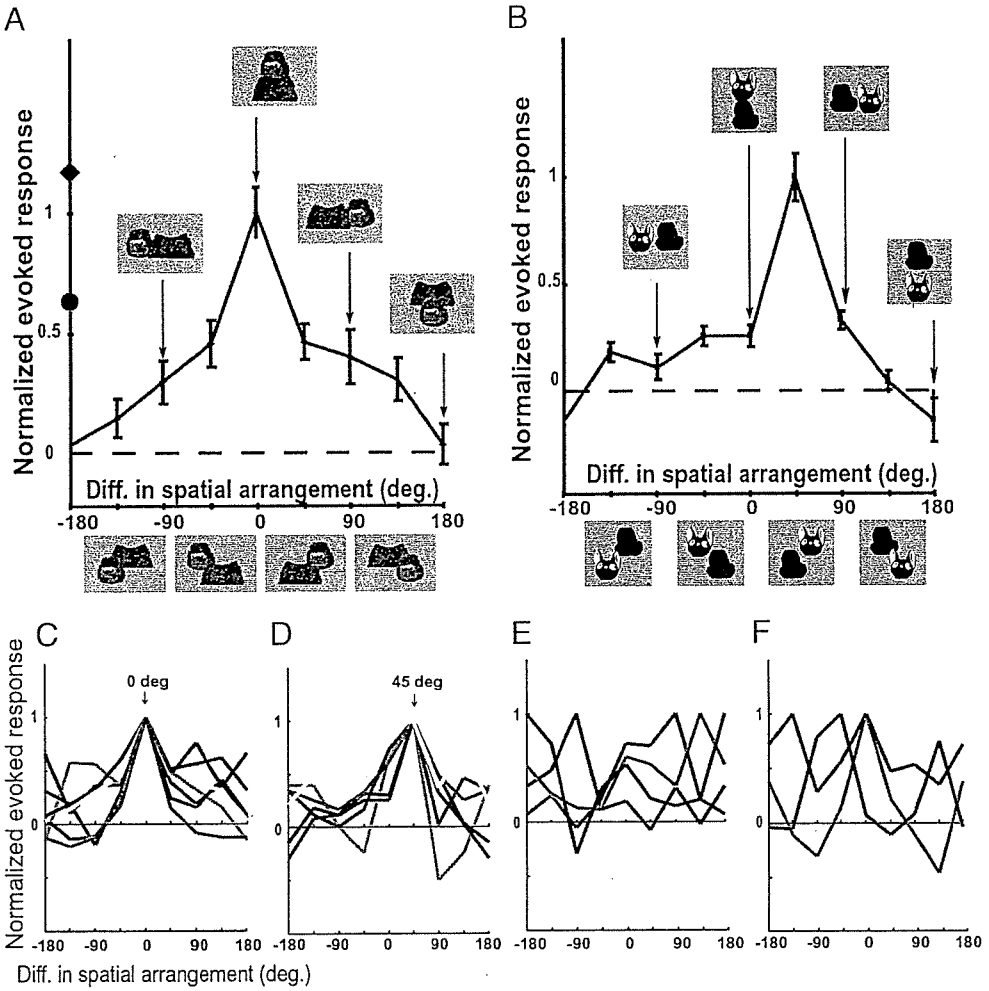



FIG. 6. Selectivity of cells in SRR spots to different spatial arrangements of the upper and the lower parts of object images. The normalized evoked responses (vertical axis) were plotted against the difference in the spatial arrangement of the parts (horizontal axis). The difference in spatial arrangement is defined by the angle between a line connecting the centers of the 2 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. *A* and *B*: responses of representative cells in spots A and B, respectively. Error bars indicate SE ($n$ = 20). One-way ANOVA confirmed that the evoked responses significantly differed depending on the spatial arrangement of the parts ($P < 0.05$). Normalized responses to the original stimuli presented at different retinotopic positions are indicated by symbols along the vertical axis in *A*. The distance between the leftmost (●) and rightmost (◆) stimuli was 12.8°. *C–F*: tuning curves for other cells in spots A–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.

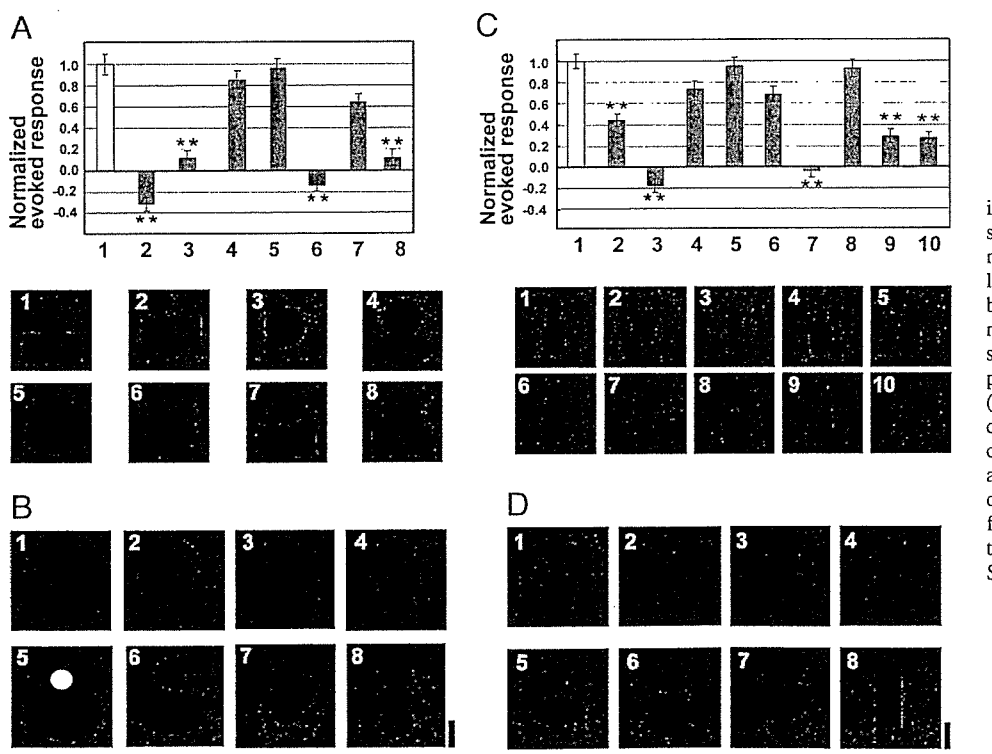



FIG. 7. Visual features critical for the cells in SRR spots. *A* and *C*: single-cell responses in spots A (*A*) and B (*C*). Evoked responses normalized to those by the critical features (stimulus 1) were plotted against the stimulus numbers. The pictures of stimuli with the stimulus number are shown below. Asterisks indicate a significant decrease in evoked responses compared with responses to the critical features (stimulus 1) ($t$-test, $P < 0.01$). Error bars indicate SE ($n$ = 20). *B* and *D*: representative critical features for different cells in spots A (*B*) and B (*D*). The stimulus was filled with black if color, luminance, and texture were not essential for activation. Only the cells with critical features 5 and 6 (*B*) were sensitive to luminance. Scale bar, 5°.

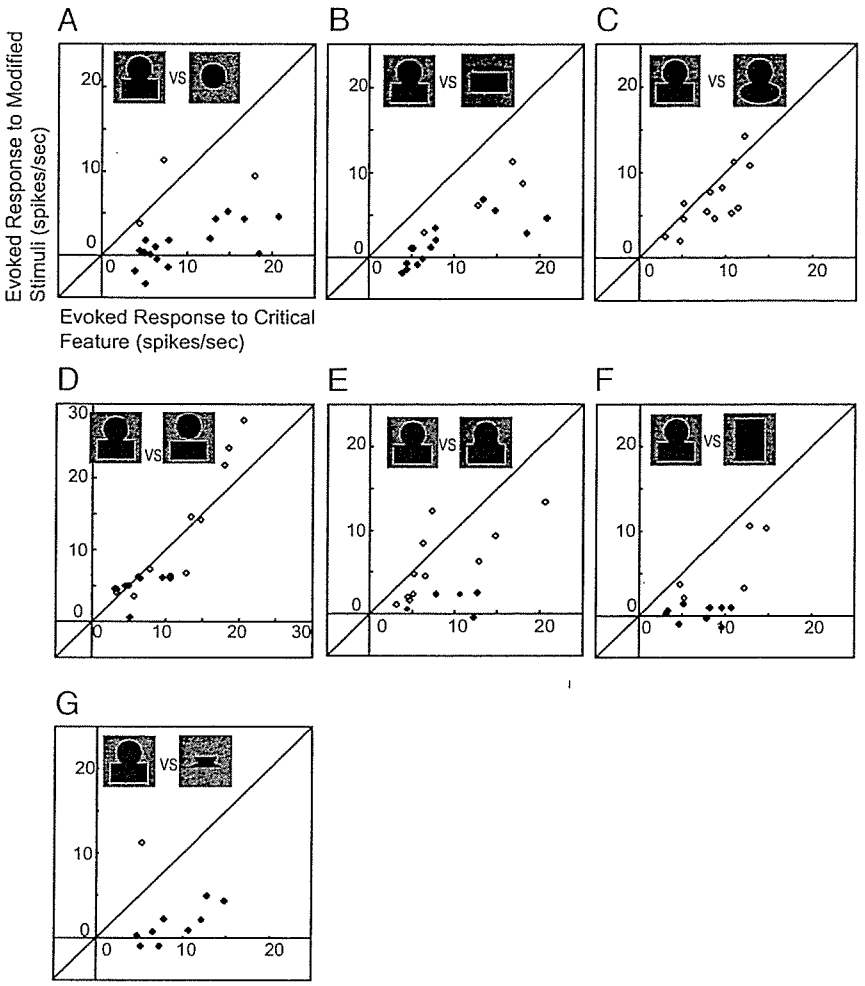



FIG. 8. Comparison between responses to the critical features and those to the modifications of the critical features. The modifications were isolated parts (A, *top*, and *B, bottom*), changes in shape of parts (*C*), inserting a gap between parts (*D*), smoothing the sharp joint (*E*), elongated shape (*F*), and isolated sharp joint (*G*). These modifications are signified by icons at the top of each graph, but these icons were not exactly the ones used in the experiment. The actual stimuli were created by modifying the critical features of individual cells. Each point represents the evoked response of single cell to its critical feature (horizontal axis) and that to modifications (vertical axis). Points below (or above) the diagonal line indicate that the response to the modification was smaller (or larger) than the response to the original critical feature. Filled symbols represent the cells that showed a significant difference in response to critical features and the modifications. These single cellular responses were obtained from spots A (red symbols) and B (blue symbols).

not respond to the combination of parts when the upper part was rotated 180° (Figs. 6 and 9). This means that, although neurons in these spots are less sensitive to local features, they still possess the capability to distinguish upper parts from lower parts by certain features existing within the parts. Because the upper and lower parts of each critical feature were different in size for most of these cells (Fig. 7, *B* and *D*), relative size may be one determining factor.

Taken together, at the level of the simplest visual features that maximally activate individual cells, we found that cells in SRR spots required two distinguishable parts for activation. For maximal activation of these cells, these two parts had to be arranged in particular spatial configurations, but it was not necessary to have particular local features embedded in parts.

## DISCUSSION

It has been shown that object images are represented by combinations of neurons representing component visual features that are less complex than object images in IT cortex (Desimone et al. 1984; Fujita 1993; Fujita et al. 1992; Ito et al. 1995; Kobatake and Tanaka 1994; Tanaka et al. 1991; Tsunoda et al. 2001; Wang et al. 1996, 1998). As for the spot indicated by the arrowhead in Fig. 4*A*, some of these neurons represent local features of object images (see also Tsunoda et al. 2001). Thus for complete reconstruction of an object image from these features, neural mechanisms that specify spatial configuration of these local features are needed. A general framework of object representation is the structural description where object images are composed of parts and the spatial relation among parts (Biederman 1987; Marr and Nishihara 1978). Although parts in this general framework and visual features represented in IT cortex are not necessarily the same, the representation of spatial configuration of elementary components is the common central issue.

Recently, Brincat and Connor examined visual responses of IT neurons with variations of two-dimensional (2D) silhouettes consisting of multiple curvatures and found that optimal features of these neurons were a combination of specific local curvatures arranged in particular positions in space (Brincat and Connor 2004). In this study, they examined the cells with 2D silhouettes but not with object images. Because representation of spatial configuration requires the cell to be insensitive to the visual attributes specific to particular parts, their results were not conclusive with respect to representation of spatial configuration of object parts.

From our findings in the present study, we suggest that neurons in area TE could represent a particular spatial arrangement of object parts based on two observations: *1*) with intrinsic signal imaging, we found activity spots that responded to a combination of two parts but not to either part shown in isolation (Fig. 4) and *2*) neurons recorded in these SRR spots were selectively activated by stimuli in which the parts were