Available online at www.sciencedirect.com
ScienceDirect
NEUROCOMPUTING
Neurocomputing 70 (2007) 1813–1818
www.elsevier.com/locate/neucom

A functional role of interaction between IT cortex and PF cortex in visual categorization task

Yoshiki Kashimori[a,b,*], Yu Ichinose[b], Kazuhisa Fujita[b]

[a]*Division of Bioinformatics, Department of Applied Physics and Chemistry, University of Electro-Communications, Chofu, Tokyo 182-8585 Japan*
[b]*Department of Information Network Sciences, Graduate School of Information Systems, University of Electro-Communications, Chofu, Tokyo 182-8585 Japan*

Available online 6 November 2006

Abstract

The ability to group visual stimuli into meaningful categories is a fundamental cognitive process. Some experiments are made to investigate the neural mechanism of visual categorization. Although experimental evidence is known that prefrontal cortex (PFC) and inferior temporal cortex (ITC) neurons sensitively respond in categorization task, little is known about the role of interaction between ITC and PFC in categorization task that gives the category boundaries in relation to behavioral consequences. To address this issue, we propose a functional model of visual system in which categorization task is achieved based on functional roles of ITC and PFC. The functional role of ITC is to represent features of object parts, based on different resolution maps in early visual system such as V1 and V4. In ITC, visual stimuli are categorized by the similarity based on the features of object parts. The PFC neurons combine the information about feature and location of object parts, and generate a working memory of the object information relevant to the categorization task. The synaptic connections between ITC and PFC are learned so as to achieve the categorization task. The feedback signals from PFC to ITC enhance the sensitivity of ITC neurons that respond to the features of object parts critical for the categorization task. In the present study, we present a neural network model, which makes categories of visual objects depending on categorization task. We investigated the neural mechanism of the categorization task of line drawings of faces used by Sigala and Logothetis (N. Sigala, N.K. Logothetis, Visual categorization shapes feature selectivity in the primate temporal cortex, Nature 415 (2002) 318–320). Using this model we show that ITC represents similarity of face images based on the information of the resolution maps in V1 and V4. We also show that PFC generates a working memory state, in which only the information of face features relevant to the categorization task are sustained. It is further shown that the feedback connections from PFC to ITC may play an important role in extracting the diagnostic features of visual images in the categorization task.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Categorization task; Feedback signal; Inferior temporal cortex; Prefrontal cortex; Neural model

1. Introduction

Visual categorization is fundamental to the behavior of higher primates. Our raw perceptions would be useless without our classification of items such as animals and food. The visual system has the ability to categorize visual stimuli, which is the ability to react similarity to stimuli even when they are physically distinct, and to react differently to stimuli that may be similar. How does the brain group stimuli into meaningful categories?

Some experiments are made to investigate the neural mechanism of visual categorization. Freedman et al.[1] examined the responses of neurons in the prefrontal cortex (PFC) of monkey trained to categorize animal forms (generated by computer) as either "doglike" or "catlike". They reported that many PFC neurons responded selectively to the different types of visual stimuli belonging to either the cats' or the dogs' category. Sigala and Logothetis [12] recorded from inferior temporal cortex (ITC) after monkey learned a categorization task, and found that

*Corresponding author. Division of Bioinformatics, Department of Applied Physics and Chemistry, University of Electro-Communications, Chofu, Tokyo 182-8585 Japan. Tel.: +81 424 43 5470; fax: +81 424 89 9748.
E-mail addresses: kashi@pc.uec.ac.jp (Y. Kashimori), makemale@yahoo.co.jp (Y. Ichinose), k-z@nerve.pc.uec.ac.jp (K. Fujita).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2006.10.067

selectivity of the ITC neurons was significantly increased to features critical for the task. The numerous reciprocal connections between PFC and ITC could allow the necessary interactions to select the best diagnostic features of stimuli [3]. However, little is known about the role of interaction between ITC and PFC in categorization task that gives the category boundaries in relation to behavioral consequences.

To address this issue, we propose a functional model of visual system in which categorization task is achieved based on functional roles of ITC and PFC. The functional role of ITC is to represent features of object parts, based on different resolution maps in early visual system such as primary visual cortex (V1) and V4. In ITC, visual stimuli are categorized by the similarity based on the features of object parts. Posterior parietal (PP) represents the location of object part to which attention is paid. The PFC neurons combine the information about feature and location of object parts, and generate a working memory of the object information relevant to the categorization task. The synaptic connections between ITC and PFC are learned so as to achieve the categorization task. The feedback signals from PFC to ITC enhance the sensitivity of ITC neurons that respond to the features of object parts critical for the categorization task, thereby enabling the visual system to perform quickly and reliably task-dependent categorization.

In the present study, we present a neural network model, which makes categories of visual objects depending on categorization task. We investigated the neural mechanism of the categorization task about line drawings of faces used by Sigala and Logothetis [12]. Using this model we show that ITC represents similarity of face images based on the information of the resolution maps in V1 and V4. We also show that PFC generates a working memory state, in which only the information of face features relevant to the categorization task are sustained. It is further shown that the feedback connections from PFC to ITC may play an important role in extracting the diagnostic features of visual images in the categorization task.

2. Model

To investigate the neural mechanism of visual categorization, we made a neural network model for a form perception pathway from retina to prefrontal cortex. The network structure of our model is illustrated in Fig. 1. The model consists of eight neural networks corresponding to the retina, lateral geniculate nucleus (LGN), V1, V4, ITC, PP, PFC, and premotor area, which are involved in ventral and dorsal pathway [9,11].

The retinal network is an input layer, on which the object image is projected. The retina has two-dimensional lattice structure that contains $N_R \times N_R$ pixel scenes. The retinal output is a gray-scale intensity of visual image received by the retinal receptors. The model of LGN consists of three different types of networks with respect to the spatial resolution of contrast detection, fine-tuned neurons with high spatial frequency (LGNF), middle-tuned neurons with middle spatial frequency (LGNM), and broad-tuned neurons with low spatial frequency (LGNB), as shown in Fig. 1. Each network receives the retinal outputs with ON center-OFF surrounding connections, and transforms the stimulus intensity of retina image into the firing rates of LGN neurons. The ON center-OFF surrounding connections were implemented by convolving the retinal outputs with a Mexican-hat-like filter. The difference in the spatial resolution of LGN maps was made by varying the shape of the Mexican-hat-like filter.

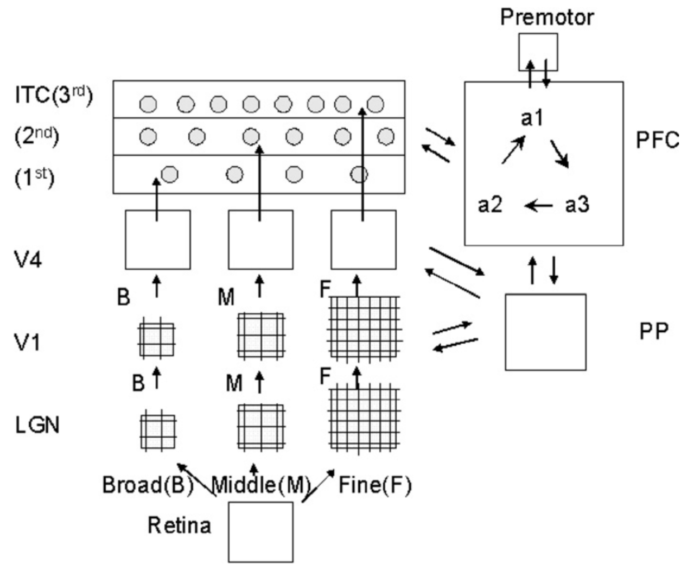

Fig. 1. The structure of our model. The model is composed of eight modules structured such that they resemble ventral and dorsal pathway of the visual cortex, retina, lateral geniculate nucleus (LGN), primary visual cortex (V1), V4, inferior temporal cortex (ITC), posterior parietal (PP), prefrontal cortex (PFC), and premotor area. a1 ~a3 mean dynamical attractors of visual working memory.

The neurons of V1 have the ability to encode elemental features of object image, such as orientation and edge of a bar. The network model of V1 consists of three different types of networks with high, middle, and low spatial resolutions, V1F, V1 M, and V1B, each of which receives the outputs of LGNF, LGNM, and LGNB, respectively. The V1X network (X = F, M, B) contains $M_X \times M_X$ hypercolumns, each of which contains L_X orientation columns. The output of V1X neuron tuned to a specific orientation was calculated using a Gabor function. The difference in the three resolution maps was implemented by varying the width and spatial frequency of the Gabor function.

The V4 network consists of three different networks with high, middle, and low spatial resolutions, V4F, V4 M, and V4B, which receive the outputs of V1F, V1 M, and V1B, respectively. The convergence of outputs of V1 neurons enables V4 neurons to respond specifically to a combination of elemental features such as a cross and triangle of the retinal images. To achieve the response ability of V4 neurons, Kohonens self-organized map model was used as a model of the V4 network.

The PP network consists of a two-dimensional array of neurons, each of which corresponds to a spatial position of each pixel of the retinal image. The PP neuron model was made based on leaky-integrator model. The functions of PP network are to represent the spatial position of a whole object and the spatial arrangement of its parts in the retinotopic coordinate system and to mediate the location of the object part to which attention is paid.

The network of ITC consists of three subnetworks, which receive the outputs of V4F, V4 M, and V4B maps, respectively, as shown in Fig. 1. It is shown that moderately complex object features are represented in posterior regions of ITC, and partial or complete object views are represented in anterior regions of ITC [2,8,10,14]. The subnetworks represent object features depending on their complexity. Each network has neurons tuned to various features of the object parts depending on the resolutions of V4 maps: The first subnetwork detects a broad outline of a whole object, the second subnetwork detects elemental figures of the object that represent elemental outlines of the object, and the third subnetwork represents the information about the fine shapes of the object parts based on the fine resolution of V4F map. Each subnetwork was made based on Kohonens self-organized map model [6]. The elemental figures in the second subnetwork may play an important role in extracting the similarity to the outlines of objects.

PFC represents the information of an object form by combining the information of the object feature encoded by the second layer of ITC and that of the location of object part represented in PP. The functional role of PFC is to generate a working memory of the object form relevant to the categorization task, based on reward signal from premotor area. The model of working memory in PFC was made based on the dynamical map model [4,5], in which memory is represented by a dynamical attractor. The network model consists of three types of neurons, M, R, and Q neurons [6]. M neuron is a main neuron, while R and Q neurons are interneurons, each of which is connected M neuron with excitatory and inhibitory synaptic connections, respectively. M neurons are interconnected to each other with excitatory and inhibitory synaptic connections. The signal propagation between M neurons has various time delays, which play an important role in the stabilization of temporal sequence of firing pattern. The model of PFC neurons was made based on leaky-integrator model. The synaptic connections between M neurons were learned by Hebbian rule depending on the timing of pre- and post-synaptic potentials, thereby enabling the PFC network to make dynamic linkage among memory attractors embedded in the PFC network.

The model of premotor area consist of neurons whose firing lead to behavior in the categorization task. The premotor neurons receive the outputs of the PFC neurons. The firing of premotor neuron relevant to the categorization task strengthens the synaptic connections between PFC and premotor area, whereas the firing irrelevant to the task suppresses the synaptic learning.

The synaptic connections between PFC neurons were made by Hebbian learning, depending on a reward signal from premotor area. They are strengthened by reward signals from premotor area when a monkey chooses correct behavior, so that only the memory relevant to the categorization task is stabilized in PFC. The connections are not strengthened due to the lack of reward signals when the monkey chooses incorrect behavior, so that the working memory in relation to the incorrect behavior is not stabilized in PFC.

The synaptic connections between PFC and ITC were also made by Hebbian learning. The retention of working memory relevant to the categorization task results in the increased connections between PFC neurons and ITC neuron encoding the information of object parts relevant to the task, leading to high sensitivity of the ITC neurons to the task relevant stimuli due to the feedback signals from PFC to ITC.

The mathematical description of our network model of visual pathway from retina to PFC is described in Refs. [6,7,13].

3. Results

3.1. Categorization task

We used the line drawings of faces used by Sigala and Logothetis [12] to investigate the neural mechanism of categorization task. The face images consist of four varying features, eye height, eye separation, nose length, and mouse height. The monkeys were trained to categorize the face stimuli depending on two diagnostic features, eye height, and eye separation. The two diagnostic features allowed separation between classes along a linear category boundary. The face stimuli were not linearly separable by using the other two, non-diagnostic features, or nose length, and mouth height. On each trial, the monkeys saw one face stimulus and then pressed one of two levers to indicate category. Thereafter, they received a reward only if they chose correct category. After the training, the monkeys were able to categorize various face stimuli based on the two diagnostic features. In our simulation, we used four training stimuli shown in Fig. 2a and test stimuli with varying four features.

3.2. Information processing of visual images in ITC and PFC

Fig. 2b shows the ability of ITC neurons encoding eye separation and eye height to categorize test stimuli of faces. In Fig. 2b, the test stimuli with varying two features were categorized by the four ITC neurons learned by the four training stimuli shown in Fig. 2a, suggesting that the ITC neurons are capable for separating the test stimuli into

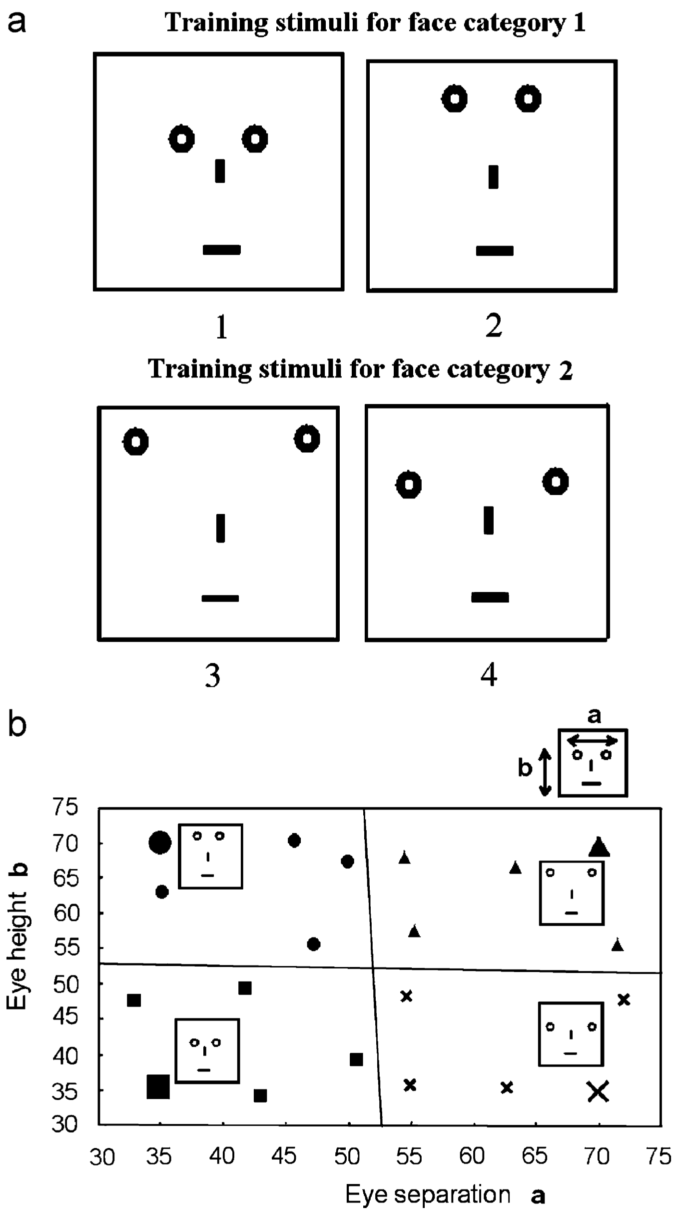

Fig. 2. (a) The training stimulus set consisted of line drawing of faces with four varying features: eye separation, eye height, nose length, and mouth height and (b) ability of ITC neurons to categorize face stimuli for two diagnostic features. The four ITC neurons were made by using the four kinds of training stimuli shown in (a), whose features are represented by four kinds of symbols (circle, square, triangle, cross). The test stimuli, represented by small symbols, are categorized by the four ITC neurons. The kind of small symbols means the symbol of ITC neuron that categorizes the test stimulus. The solid lines mean the boundary lines of the categories.

some categories, based on similarity to the features of face parts. Similarly, the ITC neurons encoding nose length and mouth height separated the test stimuli into other categories on the basis of similarity of the two features. However, the classification in the ITC is not task dependent, but is made based on the similarity of face features.

The PFC combines the information of face features and that of location of face parts to which attention is directed, and then makes memory attractors about the information. Fig. 3a shows temporal variation of the memory attractors in PFC. The information about face parts with the two diagnostic features is represented by attractors X_n ($X = \alpha$, β, γ, δ; $n = 1, 2$), in which X_1 represents the information about eye separation and eye height of the four training stimuli and X_2 represents the information about the location around eyes. The X_1 and X_2 make a pair of attractors, and the attractors X_n are dynamically linked in the PFC, as shown in Fig. 3a. As shown in Fig. 3a, the information about face parts with the diagnostic features are memorized as working memory $\alpha n \sim \delta n$ ($n = 1,2$), because the synaptic connections between PFC and premotor area are strengthened by a reward signal given by the choice of correct categorization. On the other hand, the information about face parts with non-diagnostic features are not memorized as a stable attractor, as shown by $\varepsilon 1$ and $\varepsilon 2$ in Fig. 3a, because the information of non-diagnostic features does not lead to correct categorization behavior. Thus, the PFC can retain only the information required for the categorization task, as working memory.

3.3. *A role of feedback signals from PFC to ITC in the categorization task*

The feedback connections between ITC and PFC neurons involved in the relevant task were increased during the retention of working memory in PFC, thereby causing increased responses of ITC neurons to the diagnostic features of the face images. Fig. 3b and c show the firing rates of ITC neurons encoding the diagnostic features and the non-diagnostic feature. As shown in Fig. 3b, the ITC neuron encoding the diagnostic features increases the response to face stimulus, because the neuron receives coincidently the feedforwad signal of face stimulus and the feedback signal from PFC. While, in Fig. 3c, the ITC neuron encoding the non-diagnostic feature exhibits less firing to face stimulus, because the neuron does not receive the feedback signal from PFC. As a result, the ITC and PFC neurons exhibit strong responses to the diagnostic stimulus features of the categorization task. The result may well explain the preferred responses of ITC neurons to the diagnostic face features, shown by Sigala and Logthetis [12]. The feedback signals from PFC to ITC enhance the sensitivity of ITC neurons that respond to the features of object parts relevant to the categorization task, thereby enabling the visual system to perform quickly and reliably task-dependent categorization.

4. Concluding Remarks

In the present study, we have shown that ITC represents similarity of face images based on the resolution maps of V1 and V4, and PFC generates a working memory state, in which the information of face features relevant to categorization task are sustained. We have also shown that the feedback signals from PFC to ITC may play an

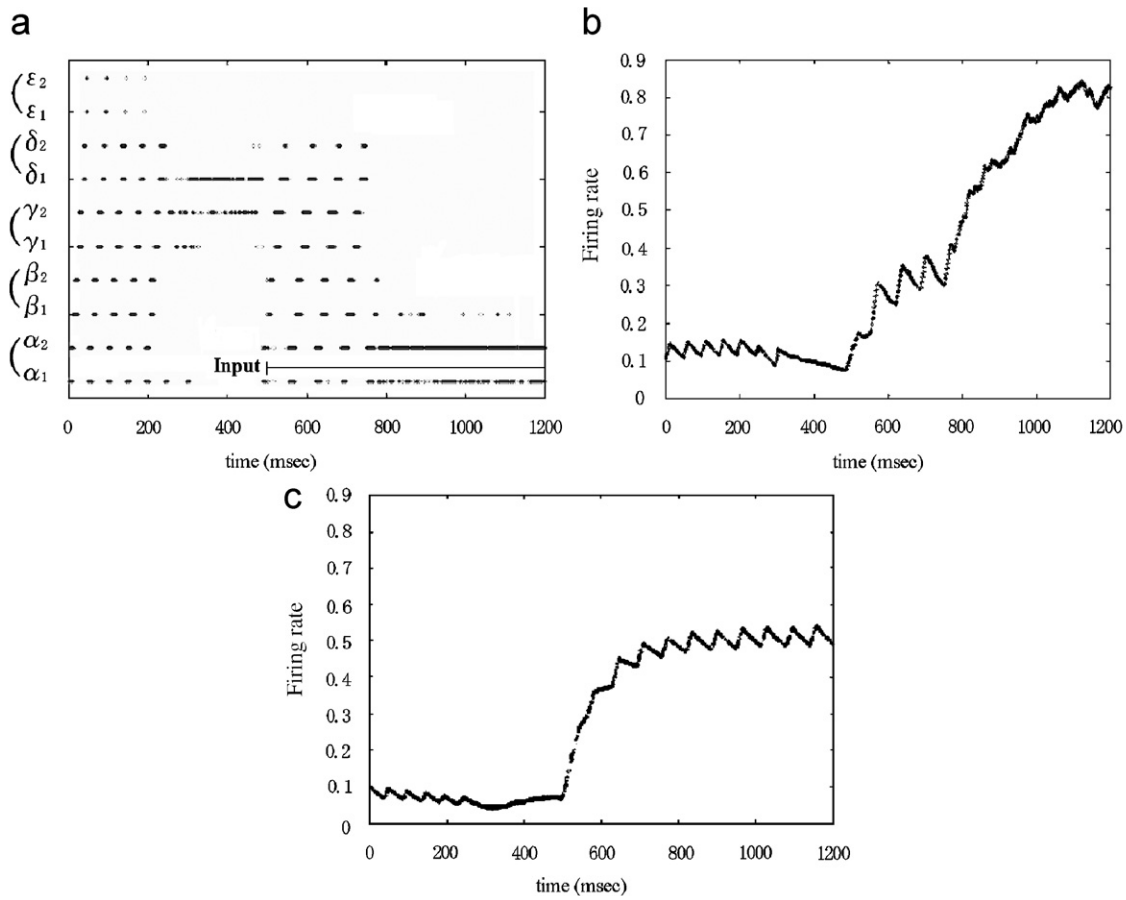

Fig. 3. (a) Temporal variation of dynamic state of the PFC network during the categorization task. Each pair attractor, X_n (X = $\alpha \sim \varepsilon$; $n = 1, 2$), consists of the feature of face part, X_1, and the location of face part to which attention is directed. The attractors representing the diagnostic features are denoted by $\alpha \sim \delta$, and the attractors representing non-diagnostic feature is denoted by ε. A mark on the row corresponding to $\alpha 1 \sim \varepsilon 2$ indicates that the network activity stays in the attractor. The visual stimulus of face 1 was applied to the retina at 500 ms. (b), (c) Firing rates of ITC neuron encoding diagnostic features (b) and that encoding non-diagnostic feature (c). The visual stimulus was applied to the retina at 500 ms.

important role in extracting the diagnostic features of visual images in the categorization task.

Our model is consistent with the experimental results of ITC and PFC in categorization tasks. It is shown by the study of Sigala and Logothetis [12] that ITC neurons show enhanced sensitivity for the diagnostic features relative to the non-diagnostic features of face images. Our model can explain the enhanced sensitivity of ITC neurons by using the feedback signals from PFC to ITC. Furthermore, the experimental study by Freedman et al. [1] demonstrated a noteworthy difference in the activity between ITC and PFC: the ITC activity quickly returned to baseline level after visual stimulus was turned off, but PFC activity remained above baseline after the stimulus was turned off. The difference in the response property between PFC and ITC can be explained by our model. In our model, ITC neuron exhibited an increased response only when it received simultaneously the feedforward signal of visual stimulus and the feedback signal from PFC. As a result, the activity of ITC neurons quickly retune to baseline after visual stimulus was turned off. On the other hand, PFC maintains its activity above baseline even after the stimulus was turned off, because the category information is being held in working memory of PFC.

We considered only the object features encoded by the second ITC layer, in which the essential features of object parts were encoded. This was because the line drawing of faces consisted of simple features. Our model is needed to extend the encoding ability of ITC so that it can extract more complex features of visual images. It is shown that ITC neurons from posterior to anterior regions respond to object features with various complexities. Further study is to investigate how the information of object features encoded by different ITC layers is used to perform visual perception.

Furthermore, our model may be useful for investigating a role of feedback signals in visual perception. The feedback signals from PFC affect not only the ITC activity but also the activity of early visual areas such as V1 and V4. Recent experimental studies show that V1 does not analyses the feedforward signals from retina, but is involved in visual perception by receiving the feedback signals from higher visual areas[15]. The present model will provide a useful insight into the role of feedback signals to V1 in visual perception.

References

[1] D.J. Freedman, M. Riesenhube, T. Poggio, E.K. Miller, Categorical representation of visual stimuli in the primate prefrontal cortex, Science 291 (2001) 312–316.
[2] I. Fujita, K. Tanaka, M. Ito, K. Cheng, Columns for visual features of objects in monkey inferior temporal cortex, Nature 360 (1992) 343–346.
[3] I. Hagiwara, Y. Miyashita, Categorizing the world: expert neurons look into key features, Nature Neurosci. 5 (2002) 90–91.
[4] O. Hoshino, S. Inoue, Y. Kashimori, T. Kambara, A hierachical dynamical map as a basic frame for cortical mapping and its application to priming, Neural Comput. 13 (2001) 1781–1810.
[5] O. Hoshino, Y. Kashimori, T. Kambara, An olfactory recognition model of spatio-temporal coding of odor quality in olfactory bulb, Biol. Cybernet. 79 (1998) 109–120.
[6] Y. Ichinose, Y. Kashimori, K. Fujita, T. Kambara, A neural model of visual system besed on mulrtiple resolution maps for categorizing visual stimuli, in: Proceedings of ICONIP2005, 2005, pp.515–520.
[7] Y. Kashimori, N. Suzuki, K. Fujita, M. Zheng, T. Kambara, A functional role of multiple spatial resolution maps in form perception along the ventral visual pathway, Neurocomputing 65–66 (2005) 219–228.
[8] N.K. Logothetis, J. Pauls, H.H. Bulthoff, T. Poggio, Shape representation in the inferior temporal cortex of monkeys, Cuur. Biol. 5 (1995) 552–563.
[9] S. Marcelija, Mathematical description of the responses of simple cortical cells, J. Opt. Soc. Am. 70 (1980) 1297–1300.
[10] E.T. Rolls, Neural organization of higher visual functions, Curr. Opin. Neurobiol. 1 (1991) 275–278.
[11] E.T. Rolls, G. Deco, Computational Neuroscience of Vision, Oxford University Press, Oxford, 2002.
[12] N. Sigala, N.K. Logothetis, Visual categorization shapes feature selectivity in the primate temporal cortex, Nature 415 (2002) 318–320.
[13] N. Suzuki, N. Hashimoto, Y. Kashimori, M. Zheng, T. Kambara, A neural model of predictive recognition in form pathway of visual cortex, Biosystems 76 (2004) 33–42.
[14] K. Tanaka, Neuronal mechanisms of object recognition, Science 262 (1993) 685–688.
[15] F. Tong, Primary visual cortex and visual awareness, Nature Rev. Neurosci. 4 (2003) 219–229.

Yoshiki Kashimori received his Ph. D. degree from Osaka City University in 1985. He is an associate professor in the Department of Applied Physics and Chemistry at University of Electro-Communications. His research interest is to clarify the neural mechanism of information processing in the electrosensory, auditory, and visual systems, based on modeling of neurons and their network. He also investigates the emergence of dynamical orders in various biological systems, based on the nonlinear dynamics.

Yu Ichinose is presently a student in the Graduate school of Information Systems at the University of Electro-Communications. His research interest is to clarify the neural mechanism of visual perception.

Kazuhisa Fujita is presently a student in the Graduate school of Information Systems at the University of Electro-Communications. His research interest is to clarify the neural mechanism of electrolocation.