# Construction of Shape Representation in Early- to Intermediate-level Visual Cortex
## --- Sparseness, Surface, and Synchronization

Graduate School of Systems and Information Engineering

University of Tsukuba

March ２０１４

Yasuhiro Hatori

# Acknowledgements

First of all, I would like to appreciate my supervisor Dr. *Ko Sakai* for his insightful suggestions and critical comments. His leading shaped my perspective of science, guiding me to pursue this work. My thanks go to the thesis committee, Drs. *Shoji Makino*, *Yukio Fukui*, *Keisuke Kameyama*, and *Hideitsu Hino* for sparing their precious time for reviewing my thesis.

Thanks to my colleagues at Computational Vision Science Laboratory. I am glad to spend much time with them. In particular, Drs. *Mitsuharu Ogiya*, *Haruka Nishimura*, and *Nobuhiko Wagatsuma* gave me a lot of suggestions. *Tatsuroh Mashita* developed the fundamental part of source code for the simulations described in Chapter 2. *Masato Sumi* performed the psychophysical experiments described in Chapter 4.

My thanks go to my friends for having a chat, going on a trip, drinking beer, and so on. All of things enrich my experience of everyday living.

Finally I would like to thank my family: my grandparents, *Iwakichi*, *Kimi*, *Haruo*, and *Tatsumi,* each teaches me many things; my uncles and aunts, especially, *Michiko* supports my daily life; my brothers and cousins, and extended family; my parents *Kazuyoshi* and *Akemi*, they understand me the most.

March 2014, *Yasuhiro Hatori*

# Abstract

Seeing is crucial to recognize the world and live in it. However, how our brain achieves the recognition has not been clarified. Construction of the neural representation of shape is a fundamental step towards shape perception and object recognition. Physiological studies have suggested that the representation of shape is established through the ventral visual pathway. Neurons in an early stage (the primary visual cortex (V1)) represent local, simple features such as orientation. Recent physiological studies have reported that neurons in V4, intermediate stage of the ventral stream, generate the representation of curvature as the subsequent representation of orientation.

Although the selectivites of neurons in the ventral stream have been reported by physiological studies, what neural mechanisms establish the selectivities has not been clarified. This thesis aims to provide computational understanding on how the visual cortices establish the neural representations of shape. Specifically, I investigated the neural mechanisms that generate the curvature representation in V4 and Medial Axis representation in V1. To clarify the neural mechanisms, I developed three distinct computational models, and carried out the simulations.

First, I developed a computational model that utilizes sparse coding, in order to investigate the coding scheme of V4. Computational studies have reported that sparse

coding generates the orientation selectivity in V1, so that the sparseness seems to be a key to clarify the coding scheme in V4. In the present study, I investigated whether V4 shares the same coding principle as in V1, because sparse coding has been widely reported in the nervous system including vision, audition, olfaction, and others. I applied component analysis with sparseness constraint to the activities of model V2 neurons in response to natural images, so as to obtain the basis functions corresponding to the receptive fields of V4 neurons. In order to investigate the dependence of computed basis functions on sparseness, I generated multiple sets of basis functions whose sparseness was systematically differed. Quantitative measurement of curvature selectivity of the basis functions showed that the bases with appropriate sparseness reproduced the characteristics of V4 neurons, suggesting the crucial role of sparseness on the construction of the curvature selectivity.

Second, I investigated the role of surface representation on the construction of the curvature selectivity. In the component analysis model, the receptive fields of model V2 neurons were the combinations of two Gabor filters: the two Gabor filters faced toward each other or aligned in a straight line with the same phase. Such configurations may yield the representation of surface, so that the surface representation appears to be essential for the generation of the curvature selectivity. To test the role of surface representation, I developed a biologically plausible model that integrates local orientations detected by V1. In the biological model, the activities were determined based on the preference for orientations and its positions of V1 neurons. I carried out the simulations with all possible combinations of the preference, so that some model neurons may generate the representation of surface, and others may not. Simulation results showed that the integration of the local orientations with surface representation

yields the curvature selectivity, suggesting the important role of surface representation.

Two distinct computational models highlight the crucial roles of sparseness and surface representation on the construction of the curvature selectivity. It is expected that the sparseness of both models should be matched, if sparseness and surface representation are essential for the construction of the selectivity. I measured lifetime sparseness of each model cell and that of each basis function. The distributions of lifetime sparseness of the biological model and the component analysis model were matched. These results suggest that sparseness and surface representation play essential role in the integration of responses in V1 and V2, in order to establish the curvature representation in V4.

Recent physiological studies have reported that V1 generates the representation of shape together with spatial information. Lee et al. have demonstrated that V1 neurons respond to Medial Axis (MA) of an object which is the set of equidistant points from nearby contours. I investigated the neural mechanism for the construction of MA representation. In this research, I took into account onset synchronization (synchronization caused by stimulus onset) of Border Ownership selective neurons that tell the direction of figure with respect to their receptive fields. The signals from BO selective neurons begin to propagate from contours simultaneously. Such signals will meet at the equidistant point from nearby contours within a short time period, yielding the representation of MA. To test the hypothesis, I developed a biologically detailed model consisting of model BO selective neurons and V1 neurons, and carried out the simulations with various stimuli including the shapes obtained from natural images. Simulation results showed that the veridical representation of MA is constructed from the synchronized signals from BO selective neurons. These results indicate that the

onset synchronization of BO selective neurons is crucial for the emergence of MA representation.

In this thesis, I developed two distinct computational models for the construction of the curvature selectivity, which utilize component analysis with sparseness constraint and spatial pooling, respectively. Both of models reproduced the characteristics of V4 neurons, and the distributions of lifetime sparseness were matched. These results suggest that the spatial pooling and sparseness plays important roles in the establishment of the representation of curvature. I also developed the biologically detailed computational model in order to investigate the neural mechanism of MA representation. The model provides the insight into the mechanism of integration of signals from BO selective neurons. Specifically, onset synchronization of BO selective neurons is crucial for the establishment of cortical representation of shape by means of MA.

# Contents

# List of figures

# List of tables

# Chapter 1. General Introduction

We recognize the world instantaneously and correctly mostly through our vision. Although we can recognize the world effortlessly, no computational algorithm can do that like humans. How does our brain recognize the world? What is a crucial step to do that? Construction of shape representation is an essential step toward shape perception and object recognition. Recent physiological studies have revealed that the representation of shape is constructed through the ventral visual stream in accordance with its hierarchical structure. For instance, the neurons in primary visual cortex (V1) represent orientation, and those in V4, subsequent region of V1, generate curvature representation. However, what neural mechanisms establish the selectivity has not been clarified. This thesis aims to provide computational understanding on how the visual cortices establish the neural representations of shape. Specifically, I investigated the neural mechanisms that generate the curvature representation in V4 and Medial Axis representation in V1. This section describes a brief summary of recent advances in the physiological research of cortical processing of shape, and gives the overview of the thesis.

## 1.1 Shape representation in the ventral stream

The construction of shape representation starts from light hitting the retina. How does our brain establish the neural representation that is useful for recognizing external world? Recent physiological studies have suggested that the cortices are organized in hierarchical manner [1], and that such organization plays important role. Specifically, the representation of shape is constructed through the ventral stream. The primary visual cortex (V1), the first area of the stream, generates the representation of orientation [2]. Successive regions generate the representation of complicated features. The neurons in V2 are selective to angles [3] and Border Ownership (BO) [4], and those in V4 are tuned to curvature and surfaces [5-8]. Inferotemporal (IT) cortex, the end of the ventral stream, encodes three dimensional shape of an object [9]. These facts indicate that the features are integrated in accordance with the hierarchy, in order to construct the useful representation for recognizing object. Brief reviews of selectivities observed in the ventral stream are given in following three sections.

## 1.2 Physiology of Medial Axis representation

Orientation selectivity is one of the import characteristics in V1. However, it is not sufficient for explaining the responses of V1 neurons. Lee et al. have been reported that V1 neurons respond to Medial Axis (MA) that is the set of symmetric points from nearby local contours [10]. In the experiment, they presented the stimuli defined by two orthogonal textures, forming an object and a background. As expected by orientation selectivity, V1 neurons responded to the texture of object when the orientation of texture corresponds to cells' preference. In addition, the neurons showed strong activities to the

center of the object, suggesting that V1 neurons represent not only orientations but also the MA of object.

## 1.3 Physiology of Border Ownership selectivity

Border Ownership (BO) selectivity is crucial to distinct the characteristics of V2/V4 neurons and those of V1 neurons [4]. BO indicates which side of the border owns the border. Zhou et al. have examined the responses to the stimuli whose local orientation and contrast projected to the classical receptive field (cRF) of cell was identical but BO directions were differed [4]. Around 20% of V1 neurons showed the selectivity to BO, and 60% of V2/V4 neurons showed the selectivity. This result suggests the crucial role of V2/V4 neurons in the BO representation. BO is essentially a local cue (e.g. [11]) that indicates the direction of figure (DOF) with respect to the cRF of the cell. In this thesis, I refer BO as the direction of figure at local point (the specific position), and DOF as the global direction of figure.

## 1.4 Physiology of Angle & Curvature selectivity

Neurons in V2 that is the successive region of V1 generate the representation of complicated features such as angle [3] in addition to BO. Ito and Komatsu have presented the stimuli composed of two lines to V2 neurons. The V2 neurons showed strong responses to a specific angle. Ito and Goda have proposed a computational model, and showed that the responses to angle could be explained by the summation of the response to each line component [12]. These results indicate that V2 neurons integrate the activities of V1 neurons that respond to orientations.

Neurons in V4 are selective to curvature that is close to the concept of angle [5-8]. Carlson et al. have examined the selectivity of V4 neurons by using genetic algorithm [8]. First, they generated stimuli randomly. Some of the stimuli drove V4 neurons, however, others did not. The next generation of stimuli was constructed by modifying the previous stimuli based on the activities of the V4 neuron. By repeating this procedure, they were able to reveal that what features in the stimuli effectively drive the neuron. The results suggested that V4 neurons respond vigorously to the specific curvature and its direction. In addition, Carlson et al. have also revealed the population characteristics of V4 neurons. They measured population activity of V4 neurons by computing the linear summation of the activities of each neuron. Population activity was biased toward acute curvature, suggesting that V4 neurons are tend to represent acute curvature.

## 1.5 Thesis overview

Recent advances in physiological studies have indicated the selectivity of neurons in the visual cortices. However, how are these selectivities constructed? What is the coding principle used in the cortices? Such questions have not been clarified. The present work aims to describe what coding scheme and neural mechanism involved in the cortical network.

Chapter 2 describes coding scheme used in V4. Specifically, I investigate whether sparse coding is crucial for the construction of the curvature selectivity. Simulation results suggest that sparse coding with input of natural images establishes the curvature selectivity. Chapter 3 examines neural mechanism for the construction of the curvature selectivity. I developed a biological plausible model and carried out the

simulations. The analysis of the model cells suggests that the representation of surface together with the integration of local orientations is essential for the establishment of the curvature selectivity. Chapter 4 presents the neural mechanism that generates the representation of MA. Simulation results suggest that the synchronized signals from BO selective neurons play crucial role in the construction of MA representation. Chapter 5 describes the summary of the thesis, and the significances of the present work. Subsequently, I discuss the directions for future work.

# Chapter 2. Sparse coding in V4

Physiological studies have reported that V4 neurons are selective to curvature and its direction. What does coding scheme account for the construction of the selectivity? Sparse coding seems to be the key to clarify this issue, because it has been reported to generate orientation selectivity in V1. The present section investigates whether sparse coding generates the selectivity in V4.

## 2.1 Introduction

Olshausen and Field have suggested computationally that sparseness plays crucial role in the construction of orientation selectivity in V1 [13]. They have assumed that images can be represented by a linear superposition of basis functions and their coefficients:

$$I_i = \sum_j c_{ij} B_j, \tag{2-1}$$

where, $I_i$ represents $i$th image, $B_j$ represents $j$th basis function, $c_{ij}$ represents coefficient of $j$th basis function to $i$th image. The bases and their coefficient were obtained by component analysis with sparseness constraint on the coefficients in terms of maximization of sparseness. Maximization of the sparseness means that relatively a small amount of absolute coefficients become large value and others do not. Considering each basis function as single neuron, their coefficients represent the

activities of each neuron to the images. Sparseness maximization corresponds to reduction of the number of active neurons. The computed basis functions showed the localized structures like Gabor functions that are similar to the receptive field structures of V1 neurons [14], suggesting that sparse coding generates the orientation selectivity of V1.

## 2.2 The proposed hypothesis

In the present study, I propose that sparseness plays a crucial role in the construction of the curvature selectivity in V4. It is plausible to consider that V4 shares the same coding scheme as V1, because physiological studies have reported that sparse coding is observed in many visual areas including V4, in addition, not limited in visual cortex [15-20]. To test the proposal, I developed a computational model, and carried out the simulations with the input of natural images. I generated the basis functions whose sparseness altered systematically, and compared their selectivity to the physiology.

## 2.3 The Model

Fig. 2-1 illustrates the computational flow of the proposed model. Although the essence of computation is identical to the model proposed by Olshausen and Field [13], some modifications have been made. First, stimuli obtained from natural images were binarized so as to focus on the information of shapes (details are described in section 2.4.1). Second, component analysis with sparse constraint was applied to the activities of model V2 neurons in response to natural images. Since V4 neurons receive ascending input from V2 [1], the correspondence between the hierarchy of the model

and that of the cortex enables me to consider the obtained basis functions as the receptive fields of V4 neurons. To compute the basis functions with sparse coefficients, I solved the optimization problem that minimizes the following cost:

$$\text{cost} = [Reconstruction\ error] + \lambda[denseness], \tag{2-2}$$

where, the first term in the right hand side of the equation indicates the error of reconstruction from the basis functions and their coefficients. *Reconstruction error* is defined as the sum of squared error between the input signals and the reconstructed signals:

$$[Reconstruction\ error] = \sum_{x,y,i}\left[A_i(x,y) - \sum_j c_{ij}B_j(x,y)\right]^2, \tag{2-3}$$

where, $(x, y)$ is coordinate, $B_j$ represents $j$ th basis function, $c_{ij}$ represents coefficient of $j$th basis function to $i$th input. $A_i$ indicates the activities of model V2 neurons. The receptive fields (RFs) of model V2 neurons were comprised of the combination of two Gabor filters, mimicking angle selectivity reported physiologically [3, 12, 21].

The second term of the right hand side of eq. 2-2 evaluates sparseness of coefficients:

$$\lambda[denseness] = \lambda\sum_{i,j} S(c_{ij}/\sigma), \tag{2-4}$$

where, $S$ is a non-linear function which poses cost in proportional to absolute values of coefficients. Such non-linearity can be realized by lots of functions, I used $S(x) = \log(1 + x^2)$ as similar to the previous study [13]. $\lambda$ and $\sigma$ in eq. 2-4 indicate positive constants which are the weight of sparseness term in the cost function and the scaling factor of coefficients, respectively. The values of $\lambda$ and $\sigma$ are described in section 2.4.2.

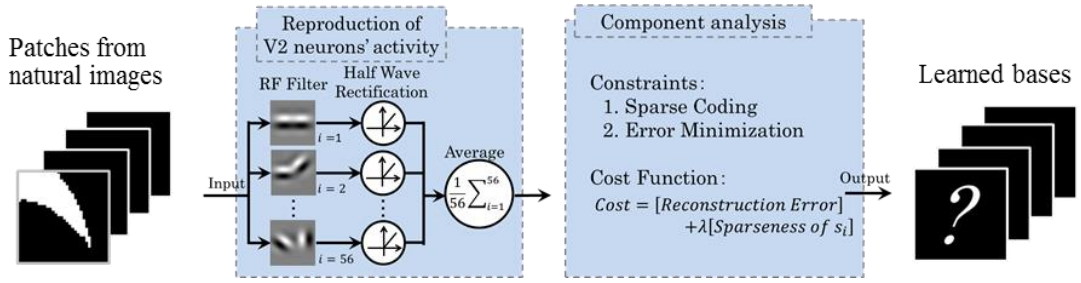


**Fig. 2-1. Schematic illustration of processing flow of the model. Component analysis with sparseness constraint is applied to the activities of model V2 neurons responding to patches obtained from natural images. The receptive fields of model V2 neurons are composed of the combination of two Gabor filters. I consider the learned bases as the receptive fields of V4 neurons.**

## 2.4 Methods

### 2.4.1 Stimuli

Input images for the model were part of natural images obtained from the Amsterdam Library of Object Images [22]. The natural images were binarized to focus solely on the information of shapes, not on other cues such as color and texture. Then, patches (33 × 33 pixel) were cut out along with the contours of the objects in order to ensure that the contours pass along the center of the patches. 80 thousands of patches were made by this procedure.

### 2.4.2 Parameters

In the present study, it is necessary to obtain the basis functions whose sparseness is altered systematically, so as to investigate the contribution of sparseness on the construction of the curvature selectivity. $\lambda$ and $\sigma$ in eq. 2-4 were changed by the factor

of 10, because these parameters determine sparseness of a code.

$$\lambda = \{2.2 \times 10^x | x = -2, -1, 0, 1, 2\}, \tag{2-5}$$

$$\sigma = \{0.316 \times 10^x | x = -2, -1, 0, 1, 2\}. \tag{2-6}$$

$\lambda = 2.2$ and $\sigma = 0.316$ were used in the previous report [13]. All possible combinations of these parameters were used in the simulations.

### 2.4.3 Population sparseness

Sparseness of a code is measured quantitatively as population sparseness [8]. Given that the coefficients ($\mathrm{C} = \{c_{i,j}\}$), I defined population sparseness ($S_P(C)$) as follow:

$$S_P(C) = \frac{1 - a}{1 - \frac{1}{n}}, \tag{2-7}$$

$$a = \langle \frac{\langle |c_{i,j}| \rangle_j^2}{\langle c_{i,j}^2 \rangle_j} \rangle_i. \tag{2-8}$$

$n$ is the number of basis functions ($n = 64$ in the present study). Notation $\langle \cdot \rangle_i$ indicates average across $i$. Population sparseness ranges from 0 to 1. For instance, $S_P(C) = 1$ means that each input signal is represented solely by one basis function (i.e. the code is sparse).

### 2.4.4 Computation of curvature selectivity

To quantify the curvature selectivity of computed basis function, I took the convolution of stimuli defined by curvature and its direction (Fig. 2-2) with the basis function:

$$R_i = I_i * B. \tag{2-9}$$

$I_i$ is the $i$th stimulus, *B* is the basis function. Notation * represents convolution. $R_i$ was passed through half-wave rectification and sigmoidal function.

To make the comparison between the model and the physiology, I computed Spike Weighted Matrix (SWM) in the same way as the physiological study [8]. A stimulus

included distinct curvatures and their directions along its contour. The activity was mapped to a matrix defined by the domain of curvature and its direction. The activity of a basis to a stimulus ($R_i$) was stored into the bins of the matrix corresponding to the curvatures and directions along the stimulus contour. The procedure was repeated for all stimuli. The activities of each bin were divided by the sample number of the bin in order to compute an average. The averages were normalized so that the value of each bin ranges between 0 and 1.

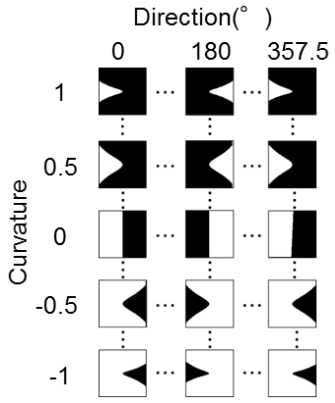


**Fig. 2-2. Test stimuli defined by curvature and its direction. Each of them has different curvature and direction.**

## 2.5 Results

To investigate whether sparseness is crucial for the construction of the curvature selectivity, I carried out the simulations so as to obtain basis functions that have distinct sparseness. I compared the selectivity of each basis functions and their population activity with those of the physiology.

### 2.5.1 Selectivity of single basis functions

Fig. 2-3 shows the examples of computed basis functions with distinct sparseness. No structure is apparent when sparseness is low (Fig. 2-3a; $S_P = 0.38$). In contrast, there are localized structures somewhat represent the specific shapes in the basis functions whose sparseness is high (Fig. 2-3b, c; b: $S_P = 0.81$, c: $S_P = 0.87$). To investigate the contribution of sparseness on the construction of the curvature selectivity, I measured curvature selectivity of these basis functions (Fig. 2-4a-c; a: $S_P = 0.38$, b: $S_P = 0.81$, c: $S_P = 0.87$). The basis functions did not reproduce the selectivity with low sparseness (Fig. 2-4a). In contrast, the basis functions whose sparseness is high (Fig. 2-4b, c)

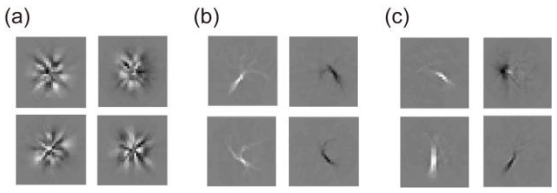


**Fig. 2-3. Examples of computed basis functions. (a) $S_P = 0.38$, (b) $S_P = 0.81$, (c) $S_P = 0.87$. No localized structure is apparent when sparseness is low (a). Localized structures are observed when sparseness is high(b, c). These structures seem to represent the specific shapes.**

showed curvature selectivity consistent with physiological studies (Fig. 2-4d; [8]). These results suggest that the curvature selectivity of the single basis functions is constructed when sparseness is high.

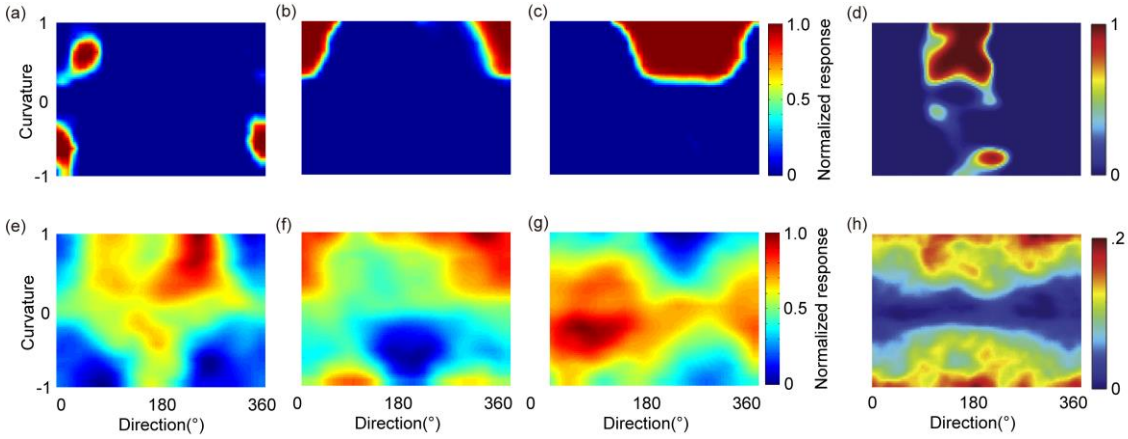


**Fig. 2-4. Activities of single basis functions and their population. The activities indicated by color (reddish means high activity) are plotted as the function of curvature and its direction. (a, e) $S_P = 0.38$, (b, f) $S_P = 0.81$, (c, g) $S_P = 0.87$. (a-c) Curvature selectivity of example single basis functions (the right bottom basis function in Fig. 2-3 a, b, and c, respectively). (d) The activity of real V4 neuron (adopted with permission from [8]). The single basis functions showed curvature selectivity when sparseness is high (b, c). (e-g) Population activities of the basis functions. (h)Population activity of actual V4 neurons (adopted with permission from [8]). The population activity showed the bias toward acute curvature when sparseness is appropriate (f).**

### 2.5.2 Acute curvature bias in population activity

In the previous section, I showed that high sparseness constructs the curvature selectivity in single basis function. Here, I addressed whether sparseness is crucial factor for the generation of population characteristics. Carlson et al. have measured population activity by computing the linear summation of the activities of single V4 neurons, and reported that population activity shows bias toward acute curvature (Fig. 2-4h; [8]). In this section, I investigated whether the population activities of the basis functions reproduce the bias.

Population activities of the basis functions are shown in Fig. 2-4e-g (e: $S_P = 0.38$, f: $S_P = 0.81$, g: $S_P = 0.87$). Population activities did not show the bias toward acute curvature when sparseness is low and high (Fig. 2-4e, g). There is a certain bias in the population activity toward acute curvature when sparseness is appropriate (Fig. 2-4f). These results indicate the crucial role of sparseness in the construction of characteristics of V4 neurons.

## 2.6 Discussion

### 2.6.1 Summary

I proposed that sparseness is crucial for the construction of the curvature selectivity in V4. To test the proposal, I developed the computational model that utilizes sparse coding, and carried out the simulations. The basis functions were computed from the activities of model V2 neurons in response to natural images. Simulation results of the single basis functions showed that the curvature selectivity is observed with high sparseness.

To investigate whether the population characteristics of the basis functions

reproduce those of the physiology, I measured population activity of the bases in the same way as the physiology. Population activity of the bases whose sparseness is appropriate showed the bias toward acute curvature consistent with the physiology. These results suggest that sparseness plays important role for the establishment of the characteristics in V4. In other words, the construction of the selectivity of V4 neurons is the consequence of the sparse coding.

### 2.6.2 Constraint on spatial structure

The implicit constrain in addition to the sparseness is the spatial structure of natural images. When the spatial structure of activities in model V2 neurons was randomized (i.e. the structure of the natural images was destroyed), the learned basis functions did not reproduce the curvature selectivity. This fact suggests that the curvature selectivity is a consequence of the sparse coding of the signals passed through the ventral stream. It would be interesting to clarify that what statistical constraint on natural images is necessary to reproduce the curvature selectivity.

### 2.6.3 Mathematical analysis of cost function[1]

It is assumed that the distribution of reconstruction error is a Gauss function, when the reconstruction error is defined by the sum of squared error [23]. For the sake of simplicity, let me consider a probabilistic model to obtain optimal coefficients ($C$) given that the specific basis functions ($B$) and input signals ($A$). Optimal coefficients are estimated by likelihood which maximizes $P(A - CB, C) = P(A|C)P(C)$. One can

[1] I would like to thank Dr. *Hideitsu Hino* for useful comments and suggestions described in the section.

assume that likelihood ($P(A|C)$) is defined by the Gaussian distribution as below:

$$P(A|C) \propto exp\left(-\frac{(A-CB)^2}{2\sigma^2}\right). \tag{2-10}$$

$\sigma$ is the standard deviation of the Gaussian. Here, the distribution of coefficients is also assumed to be a Gauss function, $P(C)$ can be described:

$$P(C) \propto exp\left(-\frac{\lambda}{2}\|C\|^2\right), \tag{2-11}$$

where, $\lambda$ is constant. Log likelihood can be computed as below. Note that the maximizing the log likelihood is equivalent to the minimizing negative log likelihood.

$$-log\big(P(A|C)P(C)\big) = -log\big(P(A|C)\big) - log\big(P(C)\big) \tag{2-12}$$

$$\propto (A-CB)^2 + \lambda\|C\|^2 \tag{2-13}$$

Eq. 2-13 is composed of (1) the sum of squared error and (2) regularization term that corresponds to the cost function used in this study. This fact indicates that minimizing the cost function (eqs. 2-2, 2-3, and 2-4) is equivalent to minimizing the negative likelihood with the assumption that the reconstruction error is a Gaussian distribution.

One important suggestion from the above discussion is that the form of cost function assumes the specific probability distribution (e.g. the sum of square assumes a Gaussian distribution). Therefore, the definition of cost function could affect the form of learned basis functions. In fact, the basis functions whose sparseness is appropriate did not yield the curvature selectivity, when reconstruction error was defined as $l_1$-norm instead of the square error ($l_2$-norm):

$$[Reconstruction\ error] = \sum_{x,y,i}\left|A_i(x,y) - \sum_j c_{ij}B_j(x,y)\right|. \tag{2-14}$$

This preliminary result implies that the assumption of probability distribution in cost function, in addition to sparseness, is important for the generation of the curvature

selectivity. Further analysis will give insights in the essence of coding scheme.

# Chapter 3. Surface constraint on the generation of curvature selectivity

In the previous chapter, I showed that sparse coding applied to the activities of V2 neurons generates the curvature selectivity. In the component analysis model, the receptive fields of model V2 neurons were the combinations of two Gabor filters: the two Gabor filters faced toward each other or aligned in a straight line with the same phase. Such configurations may yield the representation of surface, so that the surface representation appears to be essential for the generation of the curvature selectivity. Here, I investigate computationally the role of surface representation on the construction of the selectivity with a biologically plausible model.

## 3.1 Introduction

In the present section, I discuss the neural mechanism of the curvature selectivity by using a biologically plausible model that mimics cortical hierarchy in the ventral stream. The receptive fields (RFs) of neurons at the same eccentricity increase in accordance with the hierarchy (i.e. neurons in higher cortical region at the eccentricity have large RF; [24-26]). I hypothesize that V4 pools the particular combination of V1 activities with their large RF, yielding the selectivity to the curvature. Such pooling has been reported

theoretically to produce a number of selectivities [27-31]. For instance, orientation selectivity in V1 can be constructed from the spatial integration of the activities of LGN cells [27].

## 3.2 The proposed hypothesis

I propose that the integration of local orientations (the activities of V1 neurons) based on the representation of surface plays crucial role in the construction of the curvature selectivity. As stated in the General Introduction, the RFs of V1 neurons can be described by Gabor functions. Given that the pooling of V1 activities, some combination may yield the representation of surface, and others may not, depending on the combination of the phase of Gabors. For example, the positive values of the Gabors faced toward each other (Fig. 3-1), or aligned in a straight line with the same phase may generate the representation of surface. The representation of surface seems to play an important role in the construction of curvature selectivity because the shape of an object is determined by the contrast between the object and background. If the phase of Gabors is matched (i.e. surface representation may emerges), a V4 neuron pools strong activities of V1 neurons driven by a stimulus, yielding the strong activity of the V4 neuron.

**Fig. 3-1. Schematic Illustration of surface representation generated from a combination of Gabor filters. (left) An example stimulus. (right) A combination of two Gabor filters. Orange and blue indicate positive and negative values, respectively. Their phase is matched (positive values faced toward each other), yielding the representation of surface. It is expected that the curvature selectivity is constructed from the integration of orientatons when their phase is mached.**

## 3.3 The model

Fig. 3-2 illustrates the computational flow of the proposed model (details are described in Appendix A). The model takes two distinct computations. First stage corresponds to the computation of V1 neurons. The bank of 16 oriented Gabor filters detected local orientations. Detected orientations were passed through half-wave rectification and divisive normalization. All computations have been reported physiologically in V1 [32].

Second stage integrates the activities of the model V1 neurons in order to compute the activity of model V4 neuron. The model combines the two distinct types of model V1 neurons at the specific positions. I defied the activity of model V4 neuron as below:

$$R_{V4}^{(\theta_1,\theta_2,\varphi_1,\varphi_2)} = R_{V1}^{(\theta_1,\varphi_1)} + R_{V1}^{(\theta_2,\varphi_2)}, \tag{3-1}$$

where, $R_{V1}^{(\theta_i,\varphi_i)}$ is the activity of model V1 neuron, $\theta_i$ and $\varphi_i$ represent the preference of model V4 neuron for orientation and its position, respectively. The activity of model V4 neuron was passed through sigmoidal function in order to realize the nonlinear characteristics of actual neurons.

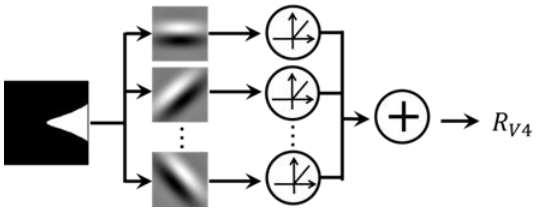


**Fig. 3-2. Schematic illustration of the computational flow of the biological model. The model extracts local orientations with 16 oriented Gabor filters. The detected orientations are passed through half-wave rectification and divisive normalization. The model pools the activities of model V1 neurons so as to obatin the activitiy of model V4 neuron.**

## 3.4 Methods

### 3.4.1 Stimuli

Stimuli used for the simulations were defined by curvature and its direction as show in Fig. 2-2. The measurement of the curvature selectivity is described in section 2.4.4.

### 3.4.2 Parameters

The activities of model neurons were determined by the two parameters: orientations ($\theta$) and their positions ($\varphi$). Both parameters were chosen from 16 orientations and 16 angular positions, respectively. I carried out the simulations with all possible combinations of those parameters, yielding 34,816 distinct model V4 neurons.

### 3.4.3 Phase analysis

I analyzed the phase of Gabor filters that are integrated by model V4 neurons in order to classify the model cells into two categories: In-phase model cell (with surface representation) and Out-phase model cell (without surface representation). I defined that surface representation emerges when the phase of two Gabor filters satisfy one of the following condition:

(i) parallel and the difference of orientations is180°,

(ii) mirror symmetry,

(iii) two Gabors located at the same position with similar orientations.

Fig. 3-3 illustrates the schematic presentation of above criteria.

### 3.4.4 Lifetime Sparseness

To compare sparseness of the biological model and that of the component analysis model, I quantified lifetime sparseness ($S_L$) of each model cell and each basis function

in the similar way as the physiology [15, 33].

$$S_L(R) = \frac{1-a}{1-\frac{1}{n}}, \tag{3-2}$$

$$a = \frac{1}{n}\left(\frac{(\sum_i r_i)^2}{\sum_i r_i^2}\right), \tag{3-3}$$

where, $r_i$ is response to the $i$th stimulus, $R$ is the set of responses($\mathrm{R} = \{r_i\}$), $n$ represents the number of the stimuli. The index ranges from 0 to 1, measuring the sharpness of the curvature selectivity. This index becomes one when model cell responds to only one stimulus. Although population sparseness and lifetime sparseness are intrinsically different measurements of sparseness, I used lifetime sparseness to compare sparseness of the biological model and that of the component analysis model. The reason is that the measurement of population sparseness of the biological model is not straightforward because population sparseness measures how much cells are used to represent images.

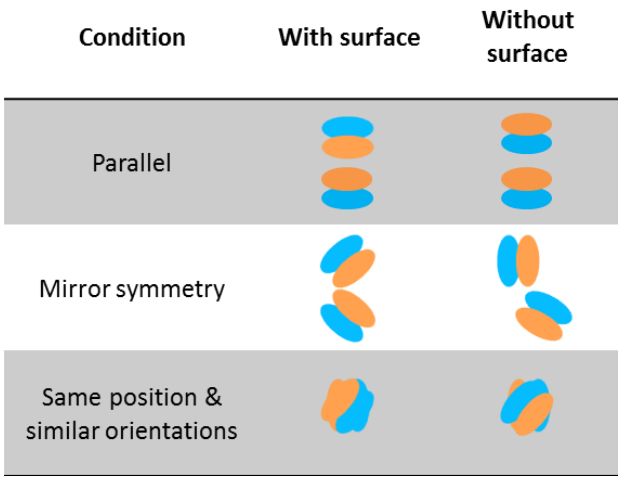


**Fig. 3-3. Schematic illustration of the criteria used in phase analysis. Orange and blue indicate positive and negative value of a Gabor function, respectively.**

## 3.5 Results

I proposed that the representation of surface generated by the phase of Gabor filters plays a crucial role in the construction of the curvature selectivity. First, I tested validity of the model in terms of reproducibility of the physiological results. Second, I investigated the role of surface representation on the construction of the selectivity. At the end, I compared lifetime sparseness of the model cells and that of the basis functions.

### 3.5.1 Selectivity of single model neurons

The activities of example model neurons are shown in Fig. 3-4a-c. The model neurons reproduced the curvature selectivity. For instance, the model neuron shown in Fig. 3-4a was selective to acute curvature pointing to the left. Each model cell showed distinct tuning for curvature and direction (Fig. 3-4a-c), reproducing the curvature selectivity reported by physiological studies (Fig. 2-4d; [8]). These results suggest that the model can reproduce the curvature selectivity at single cell level.

### 3.5.2 Population response of model neurons

I computed population activity of the model V4 cells in order to investigate whether the model reproduces the acute curvature bias. Fig. 3-4d shows population activity of the model V4 neurons. There is the bias toward acute curvature in the activity consistent with the physiological report (cf. Fig. 2-4h; [8]). This result and the results in single model cells suggest that the model is capable of reproduce the characteristics of the actual V4 neurons, indicating the validity of the model.

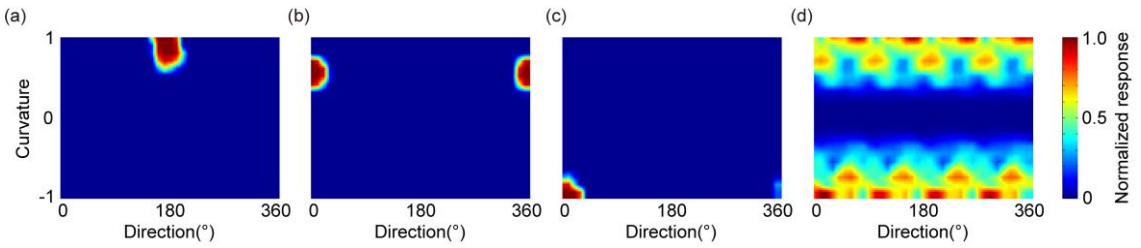


**Fig. 3-4. Simulation results of the biological model. Conventions are the same as Fig. 2-4. (a-c) Activities of three example model neurons. Each of the cells reproduces distinct tuning for curvature and direction. (d) Population activity of the model cells. It shows the bias toward acute curvature consistent with the physiology.**

### 3.5.3 **Surface constraint on the pooling of V1 activities**

In the two previous sections, I showed that the validity of the proposed model by comparing the characteristics of the model cells and those of the V4 neurons. Here, I discuss the plausibility of the proposal. I classified model cells into two distinct classes: the phase of Gabor filters is matched to yield surface representation (In-phase), or not matched (Out-phase). I investigated whether the representation of surface plays crucial role in the construction of the curvature selectivity, by comparing the selectivity of In-phase model cells and that of Out-phase model cells.

Fig. 3-5a shows scatter plots between the maximum activity of each model V4 cell and the difference of angular position of two Gabors (Difference of φ) for In-phase and Out-phase model cells (Fig. 3-5a left: In-phase, right: Out-phase). There is the tendency that Out-phase model cells do not show the strong activity for the stimuli defined by curvature and direction, suggesting that the Out-phase model cells are not selective to curvature. However, some Out-phase cells show strong activities with their phase of difference is around 22.5°. Given that the proposal, it is expected that the Out-phase model cells do not show the curvature selectivity. To clarify the proposal, further analysis was conducted to the model cells whose phase of difference is 22.5°. Fig. 3-5b shows scatter plots between maximum activities and difference of orientations (Fig. 3-5b left: In-phase, right: Out-phase). The plot suggests that the Out-phase cells show strong activities when they pool similar orientations (Difference of θ is around 0 and 337.5°). The selectivity of such cell is shown in Fig. 3-6. The cell showed selectivity for both of positive and negative curvature inconsistent with the physiology. This cell pools the same local orientations at similar positions, so that the cell operates like a line-detector. In fact the cell responds to the local orientations indicated by the grey

circle in the inset of Fig. 3-6. To summarize the simulation results: (1) In-phase model cells tend to show the strong response to curvature-defined stimuli, but Out-phase cells do not. (2) Out-phase cells that show strong response, respond to the local line, not to the curvature. These results suggest that the representation of surface is important for the construction of the curvature selectivity.

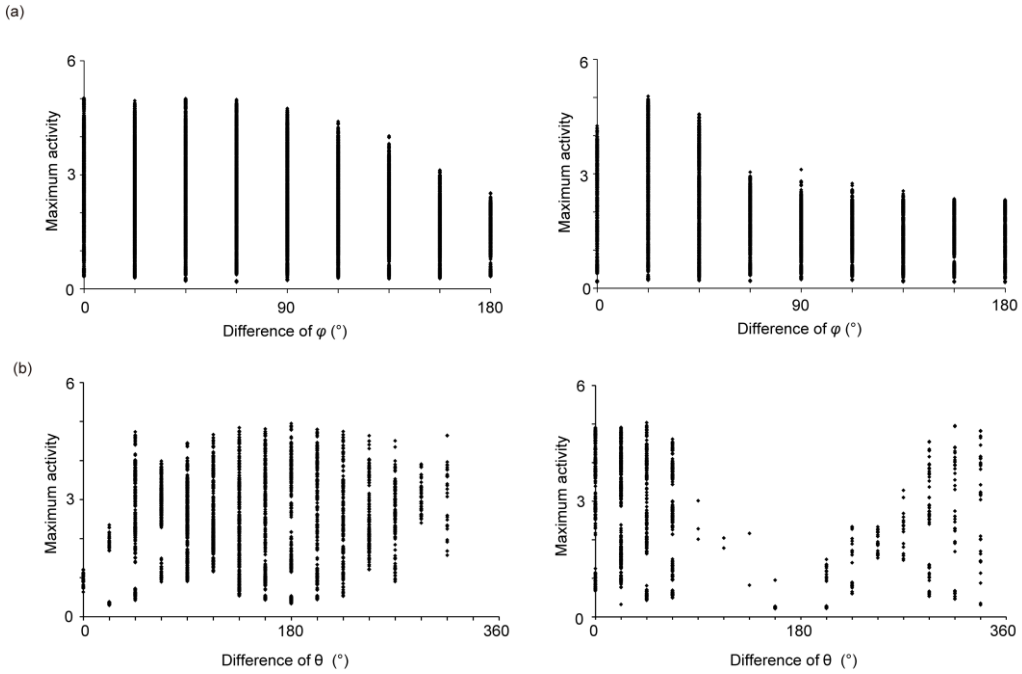


**Fig. 3-5. Contribution of surface representation on the generation of the curvature selectivity. (left column) In-phase model cells (right column)Out-phase model cells. (a) Scatter plots between the maximum response of each model cell and the difference of angular positions of Gabor filters. Although the Out-phase model cells tend to show week response to curvature defined stimuli, they respond strongly when the difference of φ is around 22.5. (b) Scatter plots between the maximum responses of each model cell and the difference of orientations of Gabor filters. Only the model cells whose difference of φ is 22.5 are plotted. The Out-phase model cell shows strong response when they pool the similar orientations at the similar positions.**

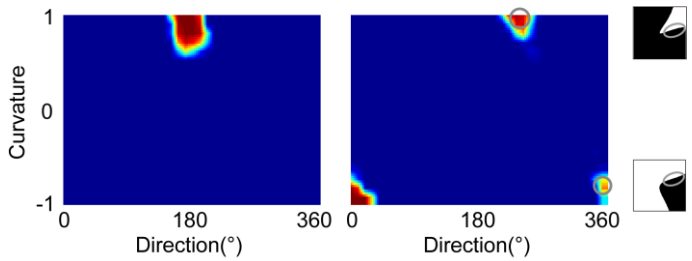


**Fig. 3-6. Comparison between selectivity of an example In-phase model cell (left) and an example Out-phase model cell (right). Although In-phase model cell reproduced the curvature selectivity consistent with the physiology, Out-phase model cell did not. It pools the similar orientations at similar positions, so that the cell detects the specific orientation denoted by gray circle in the inset.**

### 3.5.4 **Distribution of lifetime sparseness**

Two distinct computational models highlight the crucial roles of sparseness and surface representation on the construction of the curvature selectivity. It is expected that the sparseness of both models should be matched, if sparseness and surface representation are essential for the construction of the selectivity. To test this expectation, I compared lifetime sparseness of the model V4 cells to that of the basis functions. Fig. 3-7 shows the histogram of lifetime sparseness of the model V4 neurons which showed the curvature selectivity (Fig. 3-7 left) and that of the basis functions (Fig. 3-7 right; the basis functions are shown in Fig. 2-3b). The distributions are matched in terms of their mean and standard deviation (SD; 0.48 ± 0.07 for the model cells, 0.48 ± 0.08 for the basis functions; mean ± SD). The means were not significantly differed (t-test, $p > 0.60$). This result suggests that the activities of the biological model are sparse same as the component analysis model.

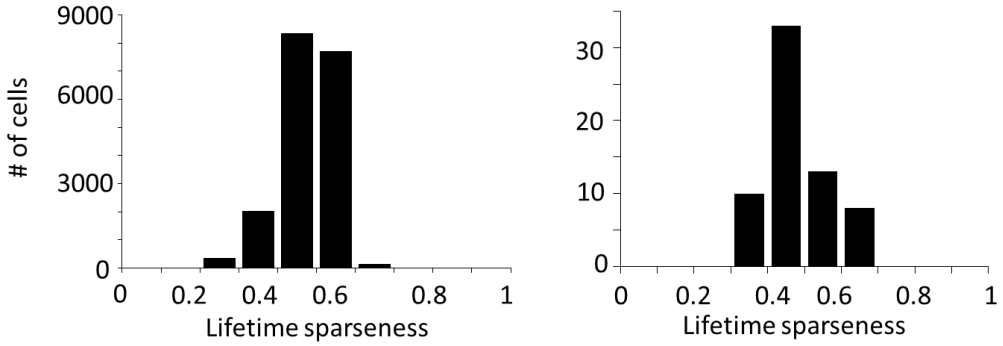


**Fig. 3-7. Distributions of lifetime sparseness of the model neurons (left) and that of the basis functions (right). Both distributions are identical (0.48 ± 0.07 for the model cells, 0.48 ± 0.08 for the basis functions; mean ± SD). The means are not significantly differed (t-test, $p > 0.60$).**

## 3.6 Discussion

### 3.6.1 Summary

I proposed that the integration of local orientations based on surface representation yields the curvature selectivity in V4. To test this proposal, I developed the biological plausible model and carried out a series of simulations. The activities of single model cells and their population were compared to those of physiology, in order to investigate the validity of the model. The model reproduced the characteristics of actual V4 neurons, suggesting the plausibility of the model.

To investigate the role of surface representation on the generation of the curvature selectivity, I classified the model cells into two categories: the phase of their Gabors generates the representation of surface, or does not. The curvature selectivity is reproduced when the representation of surface is established, suggesting the crucial role of surface representation.

Because appropriate sparseness was required for the construction of the selectivity, it is expected that the activities of the model V4 neurons with the selectivity should be sparse. To examine this idea, I compared lifetime sparseness of the model V4 cells and that of the basis functions. The distributions were identical in terms of their mean and standard deviation. This fact supports the proposal that the integration of local orientations based on the representation of surface is crucial for the construction of the curvature selectivity.

### 3.6.2 Comparison to other computational model

Some theoretical studies have proposed that spatial pooling could reproduce the curvature selectivity in V4 [30, 31]. However, their models were designed to reproduce

the physiological data. For example, Cadieu have proposed the model that obtains the optimal combination of subunits (correspond to activities of V1 neurons) fitted to the physiological data [31]. Although the model provided the curvature selectivity and position invariance reported by physiological studies, it did not clarify the essential constraint involved in the cortical network. An important advance from the present study is that the model gives insight into the fundamental neural mechanism. Specifically, the model elucidates the crucial role of surface representation on the establishment of the curvature selectivity.

# Chapter 4. Shape representation in early visual cortex

In the two precedent chapters, I focused on the coding scheme and the neural mechanism that generate the representation of curvature of contour. Here, I focus on the representation of shape obtained from the information of local contours. Recent physiological studies have reported that V1 generates the representation of shape together with spatial information by means of Medial Axis (MA). This section investigates the neural mechanism that generates MA representation in V1.

## 4.1 Introduction

The theory of MA has been developed in the purpose of efficient encoding of object shapes [34, 35]. In the theory of MA representation, the shape of an object is represented by the set of local symmetric axes and distances between the all points of the axes and the contours. Although the idea of MA is developed theoretically, recent physiological and psychophysical studies have reported the cortical representation of MA [10, 36-38]. Lee et al. have studied whether the activities of V1 neurons are modulated depending on context ([10]; see also section 1.2 ). They controlled the relative position of stimuli and the RF of V1 neuron so that the RF is located on either

contours or figure, or background. As expected by orientation selectivity, V1 neurons responded to the contour of which orientation corresponds to cells' preference. In addition, the neurons showed strong activities to the center of the object, suggesting that V1 neurons encode the MA of object not limited to contours.

V1 neurons seem to integrate the signals from BO selective neurons in V2 so as to generate the representation of MA, because (1) the size of the stimuli (4degree) exceeds the size of V1 neurons' RF (up to 1 degree at the eccentricity of 4 degree; [24]), (2) lateral connection is too slow to explain the short latency of V1 neurons that respond to MA (around 80-90ms; [10]), and (3) the modulation depends on the context that the RF of the V1 cell is onto whether figure or background. To investigate the neural mechanism of the construction of MA, I developed a biologically detailed model including V1 neurons and BO selective neurons that are mutually connected.

## 4.2 The proposed hypothesis

I propose that onset synchronization of BO selective neurons plays crucial role in the construction of MA representation. In this thesis, I referred onset synchronization as the synchronization caused by the stimulus onset. BO selective neurons signals from contours begin to propagate simultaneously, and will meet at the equidistant points from nearby contours, yielding the representation of MA. In the present model, the propagation of signals is limited to the DOF. If such directional propagation exists, the model provides the veridical representation of MA even in the presence of multiple objects. Although onset synchronization of V1 cells along contours is not sufficient for the veridical determination of MA, the synchronization among BO selective cells and the bias in signal propagation are capable of determining correct MA.

Dong et al. have reported that BO selective neurons show significant correlation [39]. Although the neural mechanism of the synchronization among BO selective neurons has not been clarified, physiological studies have reported that the spatiotemporal structures of stimulus strongly affect the synchronization of V1 cells [40, 41]. I consider that the onset synchronized activities of V1 neurons propagate to V2, yielding the synchronization in V2.

## 4.3 The model

Fig. 4-1a is the schematic illustration of model connectivity. The model is composed of two distinct layers, which correspond to V1 and V2. Each layer is comprised of model neurons described in section 4.3.1 without overlap of their RFs. The layers were mutually connected by feedforward and feedback connections. In addition to the inter-cortical connections, V1 layer has lateral connections. The characteristics of those cortical connections were determined based on the physiological evidences: the lateral connection is short (<0.5 mm) and slow (0.1 mm/ms)[42], and the feedforward/feedback connections are long (10–15 mm) and fast (3 mm/ms)[43, 44].

The computational flow of the model is described in Fig. 4-1b. The model has four distinct functional stages of computation: (i) contrast detection, (ii) determination of the DOF, (iii) integration of DOF signals, and (iv) competition by a winner-take-all. Below describes descriptions of the model cells and each functional stag. The model is implemented on the NEURON simulator [45].

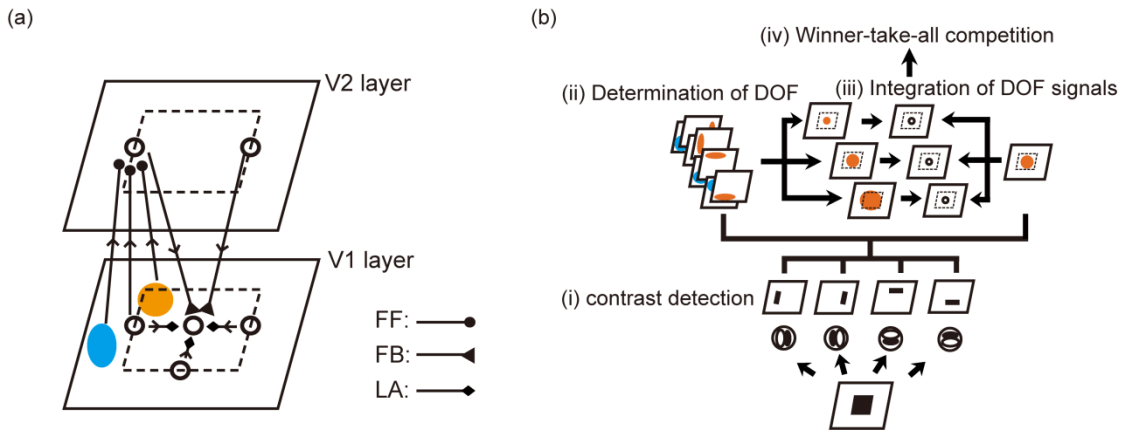


**Fig. 4-1. Schematic illustrations of the proposed model. (a) Connectivity of the model. The model is composed of two distinct layers corresponding to V1 and V2. Each layer is mutually connected in terms of feedforward/feedback connection. In addition to inter-cortical connections, V1 neurons have lateral connection. (b) Computational flow of the model. Four distinct functional stages exist: (i) contrast detection, (ii) determination of DOF, (iii) integration of DOF signals, and (iv) competition by Winner-take-all manner.**

### 4.3.1 Single model cell

Each model cell is approximated by a sphere whose diameter is 23μm. It is needed to compute precise spatiotemporal responses of the model neurons in order to investigate the manner of integration of signals from BO selective neurons. I solved the Hodgkin-Huxley equation in order to obtain the time course of membrane potential [46].

$$C_m \frac{dv}{dt} = g_{Na}(V - E_{Na}) + g_K(V - E_K) + g_l(V - E_l) + I, \quad (4\text{-}1)$$

where $C_m$ represents the membrane capacitance. $g_x$ and $E_x$ are the conductance and equilibrium potential, respectively. The subscripts $x$ indicate the type of ion channel: $Na$, $K$, and $l$ represent sodium, potassium, and other ions, respectively. $I$ is

the input current. The values of these parameters are summarized in Table 1 [47, 48]. The RF size of a model cells is 0.75°×0.75 °.

**Table 1. Parameters for each model cell.**

| Parameter | Value |
|---|---|
| $C_m$ | 1 (μF/cm$^2$) |
| $E_{Na}$ | 50 (mV) |
| $E_K$ | –77 (mV) |
| $E_l$ | –54.3 (mV) |
| $g_{Na}$ | 0.04 (S/cm$^2$) |
| $g_K$ | 0.012 (S/cm$^2$) |
| $g_l$ | 0.0001 (S/cm$^2$) |

### 4.3.2 Contrast detection

In this stage, luminance contrast is detected by four oriented Gabor filters by computing convolution of input stimulus with Gabor masks. Detected contrasts were passed through half-wave rectification and divisive normalization [32].

### 4.3.3 DOF determination

The activities of a model BO selective neuron were determined based on the luminance contrast surrounding its classical RF (cRF; details are described in Appendix B; [49]). The response of a cell is modulated by the contrast surrounding the cRF [50]. This phenomenon is called surround modulation. I assumed that a BO selective neuron has asymmetric facilitative/suppressive regions with respect to the cRF. The stimuli projected onto the surroundings of cRF modulate the activities of the model BO selective neuron in facilitative/suppressive manner depending on the configuration of surrounding regions. It has been reported that such model reproduces the characteristics of BO selective neurons [49].

### 4.3.4 Integration of DOF signal

The signals from BO selective neurons and V1 neurons responding to contours are integrated by following formula:

$$O_{\sigma}^{3}(x_2, y_2, t) = c \sum_{x,y} \left[ F_{\sigma}\left(\mathrm{t}, \mathrm{D}(x_2, y_2, x, y, V2)\right) + H\left(t, D(x_2, y_2, x, y, V1)\right) \right], \qquad (4\text{-}2)$$

where, $(x_2, y_2)$ represents the spatial location of V1 cell, $t$ is the time, $c$ is the constant for gain control. $F_{\sigma}$ and $H$ represent the feedback signals from BO selective cells and the signals from V1 neurons, respectively. $F_{\sigma}$ and $H$ are formulized as:

$$F_\sigma\big(t, D(x_2, y_2, x, y, V2)\big) = w^{\sigma}_{feedback} E\big(x, y, t - D(x_2, y_2, x, y, V2)\big), \quad (4\text{-}3)$$

$$\mathrm{H}\big(t, D(x_2, y_2, x, y, V1)\big) = w_{lateral} E\big(x, y, t - D(x_2, y_2, x, y, V1)\big). \quad (4\text{-}4)$$

$w^{\sigma}_{feedback}$ and $w_{lateral}$ is the weight of feedback and lateral connections, respectively (values are summarized in Table 2). Those weights are defined by the Gaussians whose standard deviations are 0.7, 2.1, and 3.5° for feedback, and 2.1° for lateral connection. These values are based on physiological evidence that the ranges of lateral connections are short relative to those of feedback connections [42, 43]. $D$ represents the delay of synaptic connection depending on a cortical distance. Given that the conduction velocity of each connection, $D$ can be formulized by Euclidian distance as below:

$$\mathrm{D}(x_2, y_2, x, y, L) = \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + d^2_{V1,L}}/v_L. \quad (4\text{-}5)$$

Notation $L$ indicates the origin of connection (V1 or V2). $d_{V1,L}$ represents the distance between V1 layer and $L$ layer. $L$ is 0mm if $L$ is V1, and is 30mm if $L$ is V2. $v_L$ is conduction velocity (0.1mm/ms for V1, 3mm/ms for V2; [42-44]).

**Table 2. Weights for feedback and lateral connections. Superscripts of *w* represent the spatial extent (SD) of integration fields.**

| Connection | Value |
|---|---|
| $\boldsymbol{w^{0.7}_{feedback}}$ | $0.6 * w^{2.1}_{feedback}$ |
| $\boldsymbol{w^{2.1}_{feedback}}$ | 0.008 or 0.0085 |
| $\boldsymbol{w^{3.5}_{feedback}}$ | $1.5 * w^{2.1}_{feedback}$ |
| $\boldsymbol{w_{lateral}}$ | $0.3 * w^{2.1}_{feedback}$ |

### 4.3.5 Competition by Winner-take-all

This stage determines the outputs at each spatial location by Winner-take-all manner, because three distinct integration fields yield three different activities of V1 neurons at each spatial position. Winner-take-all is the method that defines the strongest response at the position as output [51].

$$O(x,y) = max\big(S_{0.7}(x,y), S_{2.1}(x,y), S_{3.5}(x,y)\big), \tag{4-6}$$

where, $S_\sigma(x,y)$ is the number of spikes counted from $O_\sigma^3(x,y,t)$. Subscript $\sigma$ is the SD of integration field. In the present thesis, I counted a rise in membrane potential that exceed the threshold (20mv).

## 4.4 Evaluation of Reconstruction

To evaluate the accuracy of encoding by the model, I reconstructed a shape from the output (details are described in Appendix B), and calculated the reconstruction error. In briefly, I obtained the SD of integration field of V1 cells that showed the strongest response at a location. Subsequently, I placed a Gaussian with the same SD at the location. This procedure was repeated for all locations. The shape is reconstructed by the superposition of Gaussians. The reconstruction error ($Error$) was measured as the sum of squared difference between the original image ($I$) and the reconstructed image ($R$).

$$Error = \frac{\sum_{x,y}[I(x,y) - R(x,y)]^2}{\sum_{x,y}[I(x,y) + R(x,y)]^2}, \tag{4-7}$$

where $(x,y)$ represents the position. The error ranges from 0 to 1.

## 4.5 Results

I proposed that the onset synchronization of BO selective neurons is crucial for the construction of MA representation. To test this proposal, I carried out the simulations with various stimuli including natural images. First, I obtained the distribution of activities and latency of V1 neurons by using a single square, in order to test whether the model reproduce the physiological results.

One of the crucial questions is that the model could yield the veridical representation of MA for arbitrary shapes. The question is addressed by computing correlation coefficient between the representation of the model and MA obtained from a mathematical method. Another important question is whether the representation obtained from the model encodes an original shape. To investigate the accuracy of the representation, I reconstructed shape from the model output, and measured the reconstruction error.

The model assumed that the signals from BO selective neurons are propagated limited to DOF. To test the effect of the directionality in signal propagation, I carried out the simulations with the stimuli that contain multiple objects including occlusion.

Given that the hypothesis, the degree of synchronization of BO selective neurons appears to be essential factor for the construction of MA representation. I confirmed this idea by decreasing the degree of synchronization among BO selective neurons. If the degree of synchronization is crucial for the construction of MA representation, it is expected that human perception of DOF could be biased by the synchronization. Psychophysical experiment was also performed to investigate this expectation.

### 4.5.1 **Distribution of activities**

First, I carried out the simulations with a single square in order to make the comparison between the results from the model and those from the physiological studies. Fig. 4-2 shows the simulation result of a single square. The model V1 cells showed strong activity to the center of the square (Fig. 4-2a left; Distance from center = 0). The distribution is matched to that from physiological study (Fig. 4-2a right; [10]). The two dimensional distribution of the model activities are shown in Fig. 4-2b. Color indicates the response of the cell at the position. The model V1 cells whose RF is located on equidistant points from nearby contours tended to show strong response.

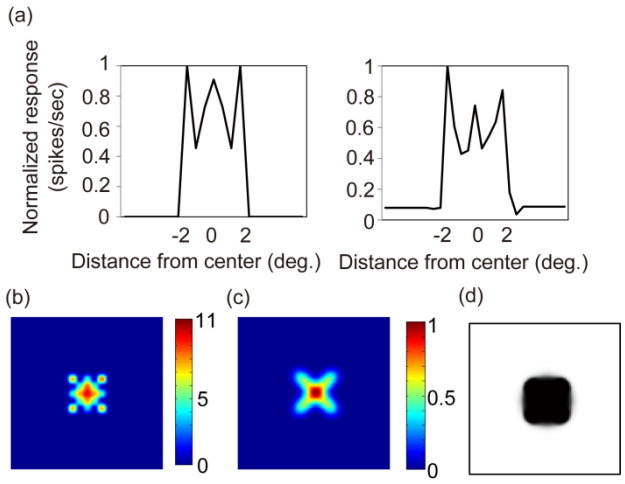


**Fig. 4-2. Simulation result for a single square. (a) The activities of the model V1 neurons whose receptive fields located onto the horizontal midline of the square (left: the model, right: the physiology [10]). Horizontal axis represents distance from the center. 0 and±2 indicate the center of the square and its contours, respectively. (b) The two dimensional distribution of the model responses. The activities were plotted as color (reddish indicates strong response). It should be noted that the responses to the contours are not shown. (c) MA representation obtained from a mathematical method. The correlation coefficient between the model output (b) and the MA from mathematics (c) was 0.91. (d) The reconstructed shape from the model output (b). Reconstruction error was around 0.03.**

### 4.5.2 **Latency of the model cells**

I made quantitative comparison of temporal characteristics between the model V1 neurons and the physiology [52]. Fig. 4-3 shows onset latency of the model (Fig. 4-3 black) and the physiology (Fig. 4-3 grey; [52]). Edge and Axis represent latency of the model V1 cells that respond to the contours and the center of the square, respectively. Although the absolute latency of the model and the physiology is differed (65.3 ms vs. 74 ms for the edge; 86.1 ms vs. 96 ms for the axis), their relative difference between Edge and Axis is similar (Fig. 4-3 diff.; 20.8 ms vs. 22 ms), suggesting that the model reflects the essence of processing flow in the cortical network.

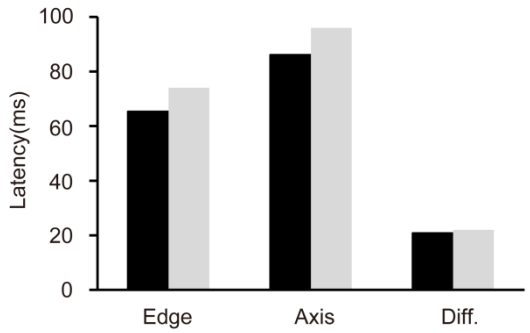


**Fig. 4-3. The latency of the model (black) and the physiology (grey; adapted from [52]). *Edge* and *Axis* represent latency of V1 cells whose RF is located along the contours and the center of the square, respectively. *Diff.* is the relative difference of latency between Edge and Axis.**

### 4.5.3 MA representation for arbitrary shapes

One of the crucial questions is that the model could yield the veridical representation of MA for arbitrary shapes. To investigate the question, I compared the output of the model and MA computed from a mathematical method by measuring correlation coefficient. The MA of the square obtained from mathematics is shown in Fig. 4-2c (using the 2-D Medial Axis Computation package of MATLAB). The correlation coefficient between the output of the model (Fig. 4-2b) and the MA obtained from the mathematical method (Fig. 4-2c) was 0.91, suggesting that the model generates the MA representation of the square.

If the model constructs the MA representation of the squares, it is of great interest to investigate whether the original shape is reconstructed from the representation. I obtained a shape from the MA representation of the model (see Fig. 4-2c), and evaluated quantitatively by measuring the reconstruction error. The reconstructed shape is shown in Fig. 4-2d. The error of reconstruction was around 3% of the maximum, indicating that the representation of the model in fact encodes the original shape.

The simulation results for the square showed that the model produce the veridical representation of MA. I also investigated whether the model yields MA representation for arbitrary shapes. I performed the simulations with the shapes obtained from natural images. The simulation results for three examples (an L-shaped tree, a stone, and a bear) are show in Fig. 4-4. Fig. 4-4a-e show the results for the L-shaped tree. The correlation coefficient between the model output and MA from mathematical method was 0.65, and the reconstruction error was 0.17. Accuracy of the representations was the same in other two shapes (the correlation coefficient and the

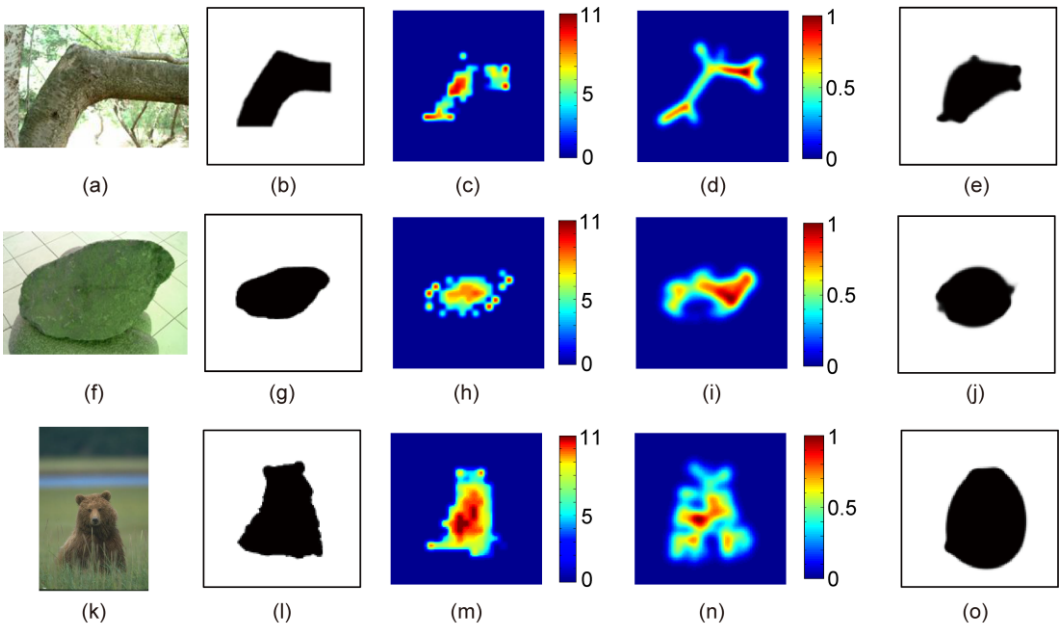


**Fig. 4-4. Simulation results for an L-shaped tree branch (a-e), a rounded stone (f-j), and a bear (k-o). The conventions used for the model outputs, the MA from the math, and the reconstructed images are the same as those described in Fig. 4-2c-e. (a, f, k) Natural images of the L-shaped tree branch, the rounded stone, and the bear. The image of the bear is obtained from the Berkeley Segmentation Dataset [53]. (b, g, l) The binary stimuli. (c, h, m) Model outputs. (d, i, n) The MA computed from the mathematical method. (e, j, o) The reconstructed shapess from c, h, and m, respectively. (a-e) The correlation coefficient between (c) and (d) was 0.65. The reconstruction error was 0.17. (f-j) The correlation coefficient between (h) and (i) was 0.76 and the reconstruction error was 0.14. (k-o) The correlation coefficient between (m) and (n) was 0.78, and the reconstruction error was 0.15.**

reconstruction error for the stone were 0.76 and 0.14, for the bear were 0.78 and 0.15, respectively). All simulation results are summarized in Table 3. The mean of correlation coefficient and reconstruction error were 0.74 and 0.16, respectively, suggesting that

the model yields the correct MA representation of arbitrary shapes.

**Table 3. The correlation with mathematical method and the reconstruction error.**

| Shape | Correlation coefficient | Error of reconstruction |
|---|---|---|
| **Square** | 0.91 | 0.028 |
| **L-shaped tree branch** | 0.65 | 0.17 |
| **Rounded stone** | 0.76 | 0.14 |
| **Bear** | 0.78 | 0.15 |
| **Two separated squares** | 0.89 | 0.028 |
| **Overlapping square (occluding)** | 0.75 | 0.082 |
| **Overlapped square (occluded)** | 0.69 | 0.038 |
| **Rectangle** | 0.67 | 0.14 |
| **Triangle** | 0.60 | 0.32 |
| **U-shaped object** | 0.73 | 0.46 |
| **Mean** | 0.74 | 0.16 |

### 4.5.4 **Directionality of signal propagation**

In the model, the propagation of signals from BO selective neurons is limited to the direction of figure. To investigate the contribution of such directionality on the emergence of MA representation, I carried out the simulations with separated two squares (Fig. 4-5a, b) and overlapping two squares (Fig. 4-5c, d). If the propagation is limited to the direction of figure, no spurious MA emerges the outside of objects. The simulation results confirm this expectation. The model produced no response to the background even if it was surrounded by multiple contours. These results suggest that the directionality in the propagation of BO signals is capable of yielding the correct MA of objects.

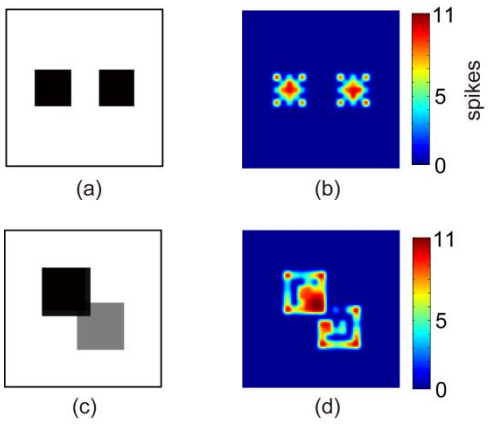


**Fig. 4-5. Contribution of the directionality in signal propagation. (a, c) Input stimuli (a: two separated squares, c: overlapping squares). (b, d) The output of the model for (a) and (c), respectively. Conventions are the same as Fig. 4-2c. No spurious MA outside of the objects emerges.**

### 4.5.5 **Effect of the degree of synchronization**

In the previous sections, I investigated the validity of the proposal that the MA representation of shape is constructed by the onset synchonization of BO selective neurons. If the synchronization of BO selective neurons is crucial, the formation of MA representation could be affected by the decrease of the degree of the synchronization. The simulations were carried out by using ambiguous figures in which two objects (regions) share the border (Fig. 4-6a). Three stimuli were the part of natural images that were taken from the Berkeley Segmentation Dataset (BSD; [53]). Onset of the two figures was diffrentiated in order to manipulate the degree of synchronization. The onset of a particular portion (90% or 60%) of the border was equated to surrounding contours of either side of the border (e.g., on the left), and the rest to those of the opposite side (on the right). Fig. 4-6 shows the simulation result for an example stimulus with 90% synchronization condition. BO selective neurons responded to the border were synchronized to those responded to either side (left or right) as shown in Fig. 4-6b. When most of the BO selective neurons synchronized to the left, the responses were biased toward the left side of the border (Fig. 4-6c left). The bias was flipped when the BO selective neurons were synchronized to the alternative side (Fig. 4-6c right). In order to evaluate the bias, I summed the activities of V1 neurons within the left and the right objects for the left- and right-synchronized conditions (Fig. 4-7a). Fig. 4-7a shows the bias toward the synchronized side. This tendency can be seen in all stimuli and both of the degree of synchronization (60% and 90%; Fig. 4-7b). Although the bias was observed in 60% synchronization condition, it was smaller than in 90% condition (Fig. 4-7b right). These results suggest that the formation of shape representation is biased toward the highly synchronized side. In addition, the bias depends on the degree of

synchronization, supporting the hypothesis.

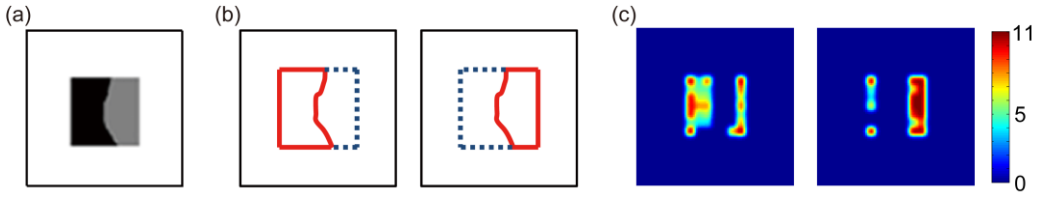


**Fig. 4-6. Simulation results for ambiguous figures. (a) An example stimulus used for the simulation. The stimuli were the part of natural images taken from the BSD [53] and filled with black and grey. (b) Schematic illustrations of synchronization condition. A portion (90% or 60%) of BO selective cells responded to the border between the ambiguous figures were synchronized with those responded to the surrounding contours of an either side of the border (the left or right; highly synchronized side is denoted by red lines), and the rest (10% or 40%) of BO-selective cells were synchronized with those responded to the opposite side of the border (denoted by blue dotted lines). (c) Output of the model when BO selective cells were synchronized (90%) with the left (left panel) and right (right panel) sides of the stimulus are shown.**

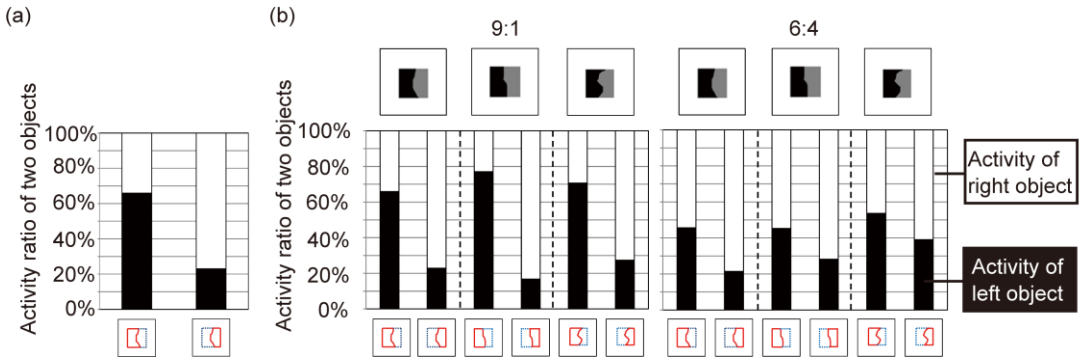


**Fig. 4-7. Quantitative analysis of the bias. (a) The summed activities of model cells within the left (black) and right (white) regions shown in Fig. 4-6c are plotted separetely in accordance with the direction of synchronization (small icons at the bottom of the plot indicate highly synchronized side). The bias toward the direction of highly synchronized side was observed. (b) The simulation results for the three stimuli. The shape of stimuli are shown at the top of the panel. The degrees of synchronization were 9:1 and 6:4 for the left and right panels, respectively. The bias was observed in all stimuli used in the simulations. However, the bias is smaller in 6:4 condition.**

## 4.6 Psychophysical experiment

The simulation results showed that synchronous presentation of a stimulus can facilitate the activity of V1 cells that respond to the MA of an object. If the synchronization facilitates the formation of shape representation, there is the possibility that a highly synchronized side tends to be perceived as figure. This phenomenon can be interpreted as the common fate, one of the Gestalt factors that are the law of grouping [54-56]. I performed a psychophysical experiment in order to investigate whether asynchronous presentation of contour elements affects the determination of DOF of an ambiguous figure. The shapes of ambiguous figures used for the psychophysical experiment were the patches of natural images as similar to the simulations. However, the contour of the stimuli consisted of blinking dots. In the experiment, the degree of synchronization is determined by the number of dots that shares the blinking cycle.

### 4.6.1 Experimental procedure

The experimental procedure is shown in Fig. 4-8. Stimuli were presented on a liquid crystal display (Mitsubishi Diamondcrysta RDT 197S; response time 5 ms; refresh rate 70 Hz). A red fixation point (0.2 × 0.2°) was shown at the center of the display with a random mask for 1500 ms. The position of fixation point was adjusted to the eye level of individual participants. After the presentation of the fixation point, a test stimulus (6.3 × 6.3°) was projected on a gray background (81.85 cd/$m^2$) for 860 ms. Test stimuli (illustrated in Fig. 4-9) was made of blinking dots whose luminance alternated between gray and either black or white (315.8 and 0.316 cd/$m^2$, respectively). The dots were placed on the border of two regions (border dots), the surrounding contours of the border that shape square (outline dots), and elsewhere (noise dots). The outline dots

formed a 4 × 4° square. The size of the dots was 0.03 × 0.03°. The spaces between dots were 0.03°.The border and outline dots were aligned in two lines (Fig. 4-9b). Only half of the border and outline dots were displayed at every moment, in order to avoid the perception of a solid contour (Fig. 4-9b). The frequency of the blinking was 7 Hz, but phases were differentiated. The degree of synchronization was defined by the number of dots that blinked with the same phase. For instance, in 90% synchronization condition, nine out of 10 dots blinked with the same phase. Outline dots placed on either the left or right of the border were more synchronized (60% or 90%) with the border dots, and the alternative outline dots were less synchronized (40% or 10%) with the border dots. The phase of each dot was randomly chosen at each presentation. The noise dots were scattered randomly but not overlapped with the border and outline dots. The

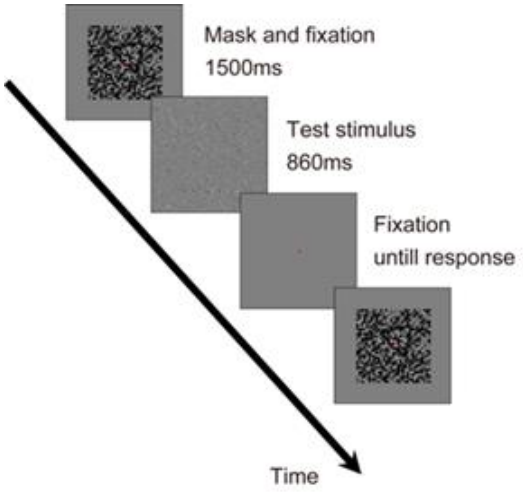


**Fig. 4-8. Schematic illustration of the experimental paradigm. A red fixation dot is presented at the center of the display with a random mask for 1500ms. Subsequently, the test stimulus is presented for 860ms. Subjects are asked to answer the direction of figure at the fixation point by two alternative forced choice. Untill subjects answered, the red fixation dot is presented.**

luminance of the noise dots was altered every 43 ms with the probability of 50%. A half of the dots were set as white and the other half were set as black, so that the mean luminance is identical at every moment. The presentation order of the stimuli and the conditions were randomly chosen. The task of participants was to answer the direction of the figure at the fixation point using two alternative forced choice task (left or right). The correct answer was not given to the participants. Five participants performed the experiment. All of them were naïve to this experiment and had normal or corrected-to-normal vision.

### 4.6.2 **Results**

I performed the experiments using following conditions which is corresponding to the simulations: the ratio of synchronization was 6:4 and 9:1, the direction of highly synchronized side was left or right. Fig. 4-10 shows the perceived DOF for each

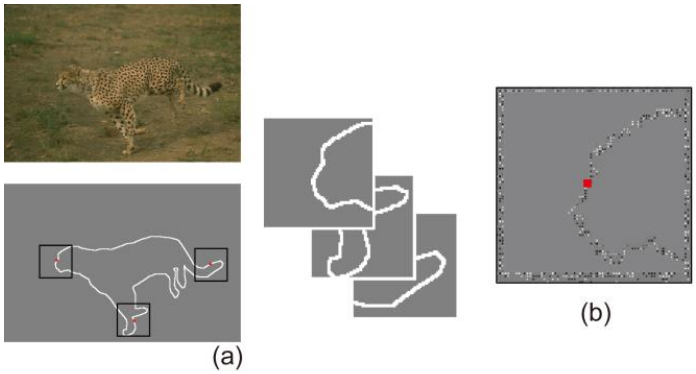


**Fig. 4-9. Configurations of a test stimulus. (a) A natural image obtained from the BSD (top left; [53]). The object contour detected from the original image (bottom left). Small patches cut out along the contour (right; positions are denoted in the bottom left image). (b) An example of a test stimulus. Border and outline dots were aligned in two lines. Note that noise dots are not shown here.**

condition. The participants tended to perceive the region as figure in the direction of the highly synchronized side compared with the opposite direction (pairwise t-test, $P < 0.01$ for both synchronization conditions). This perceptual bias increased slightly when the ratio of synchronization increased from 6:4 to 9:1. A three-way ANOVA with factors of synchronization, direction of highly synchronized side, and participants showed significance for the three main factors ($P < 0.01$) without interaction. This result indicates that the perception of figure depends on the degree of synchronization. Although the magnitude of the bias is smaller than that obtained from the simulations, the bias showed same tendency. This result supports the hypothesis that the onset synchronization of BO-selective cells is crucial for the construction of MA representation.

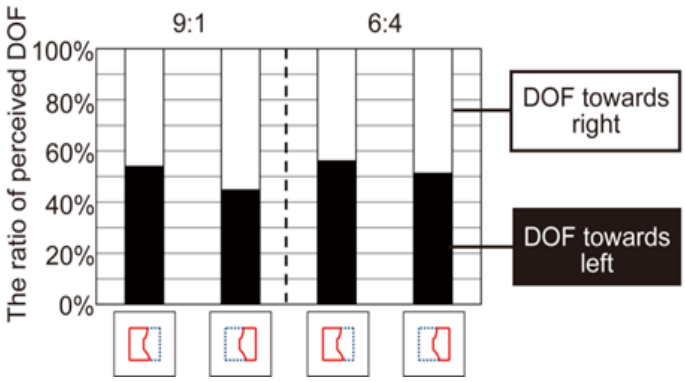


**Fig. 4-10. The perception of DOF in human observers. The results obtained for 9:1 and 6:4 synchronization ratios are plotted on on the left and right of the panel, respectively. The icons placed at the bottom of the graphs represent the highly synchronized side of the stimuli with the same convention used in Fig. 4-6. Black and white bars represent the ratio of the perceived DOF in left and right, respectively. In both the synchronization conditions, the participant tended to perceive as a figure in the highly synchronized side ($P < 0.01$; pairwise t-test). A three-way ANOVA with factors of synchronization, direction of highly synchronized side, and participants showed significance for the three main factors ($P < 0.01$) without interaction.**

## 4.7 Discussion

### 4.7.1 Summary

Here, I investigated computationally that the onset synchronization of BO selective neurons is crucial for the construction of MA representation. Synchronized signals from BO selective neurons generate strong responses at the equidistant points from nearby contours. To clarify the validity of the model, I compared the spatiotemporal characteristics of model V1 neurons and those of actual V1 neurons. The distribution of model activities matched with that of V1 neurons. The latency of the model V1 neurons was identical to that of the physiology, suggesting that the model captures the essence of processing in the cortical network.

It would be interesting to investigate whether the model yields the veridical representation of MA. I measured correlation coefficient between the representation from the model and MA computed by the mathematical method. The mean of correlation coefficient is 0.74, indicating that the model establishes the MA representations of arbitrary shapes. Another crucial question is whether the shape is reconstructed from the MA representation of the model. To address the question, I obtained shape from the MA, and compared it with original shape by measuring reconstruction error. The mean error of reconstruction is 16% of its maximum, indicating that the model yields the representation of shapes.

I investigated the contribution of directionality in the signal propagation on the construction of MA representation. Given that the directional propagation of signals from BO selective neurons, it is expected that no spurious MA outside of figure regions emerges. I carried out the simulations with two separated squares and overlapping squares. Both of the simulation results showed that the MA is established within the

objects, suggesting the crucial role of the directional propagation of signals from BO selective neurons.

If the synchronization of BO selective neurons is important for the formation of MA representation, the degree of synchronization of BO selective neurons seems to be the essential factor for the generation of shape representation. In order to investigate this idea, I carried out the simulations with ambiguous figures in which the degree of synchronization was manipulated. The degree of synchronization was controlled by differentiating the onset timing of two figures. The simulation results showed that the model responses are biased toward highly synchronized side depending on the degree of synchronization, suggesting the important role of onset synchronization among BO selective neurons. If synchronization affects the formation of shape representation, it is expected that human perception of DOF shows the bias toward highly synchronized side. To test this expectation, I performed psychophysical experiments in which the degree of synchronization of contours is controlled in the same way as the simulation. The perception of DOF was biased toward highly synchronized side consistent with the simulation results. These results support the proposal that the onset synchronization of BO selective neurons is crucial for the construction of representation of shape in terms of MA.

### 4.7.2 Comparison between another computational model

It has been reported that MA responses can be established solely by lateral connections within V1 [57]. However, it failed to reproduce the short latency of V1 cells to MA response. Physiological study has reported that the latency of V1 cells responded to the MA is around 80–90 ms after stimulus onset [10]. The lateral connection is not sufficient

to establish the response of which latency is 80–90 ms, because the conduction velocity of the lateral connection is slow (0.1 mm/ms). The conduction delay through lateral connections between the contour and the center of a square is 60-80 ms, given that the cortical distance between the contour and the center of the square is approximately 6–8 mm. The distance is estimated from following physiological evidences and experimental setting: (1) the eccentricity of the recording site is 3-4° [10] and (2) the cortical magnification factor at the eccentricity is 4.16 mm/° [58], and (3) the size between the center of the square and the contour is 2° [10]. Since conduction velocity of inter-cortical (feedforward/feedback) connections are 10 times faster than that of lateral connections (3 mm/ms), the latency of V1 cells responding to MA (80-90 ms) can be established by inter-cortical connections as shown in Fig. 4-3.

### 4.7.3 Source of synchronization

An alternative source for the generation of the synchronization among BO selective neurons is feedback signals from higher cortical areas such as V4. Synchronization by feedback signals may provide the similar results. An advantage of the present study is that the model reproduces the physiological results without the feedback from higher cortical regions such as V4. Physiological studies have reported that the timing of stimulation could affect the degree of synchronization of V1 neurons [40, 41]. For instance, Zhou et al. have reported that the synchronization of V1 cells decreases as the function of destruction of contour continuity [41]. Taking into account the cortical hierarchy, the synchronized activities of V1 neurons are propagated to V2 neurons. This fact supports the proposal that stimulus onset causes the synchronization among BO selective cells in V2.

# Chapter 5. General Discussion

In the present study, I investigated the neural mechanisms and coding scheme in different cortical layers with three computational models. Each model reflects the different aspects of cortical processing. This chapter provides the summary of simulation results of each model, the contributions of the present study, and the discussion aimed for future studies.

## 5.1 Summary of the thesis

### 5.1.1 Sparse coding in V4

I investigated the coding scheme that accounts for the emergence of curvature selectivity in V4. I applied component analysis with sparseness constraint to the activities of model V2 neurons in response to natural images, so as to obtain basis functions corresponding to the receptive fields of V4 neurons. I investigated dependence of sparseness on the construction of the curvature selectivity. The obtained basis functions with appropriate sparseness had localized structures, and reproduced the characteristics of V4 neurons. These results suggest that sparseness is crucial for the construction of the curvature selectivity. In other words, the curvature selectivity emerges as the consequence of sparse coding together with the input of natural

images.

### 5.1.2 Surface representation for the curvature selectivity

To clarify the neural mechanism that generates the curvature selectivity, I developed the biologically plausible model which integrates the local orientations detected by V1. I carried out the simulations, and investigated what constraint in the integration is necessary for the generation of the selectivity. Simulation results suggest that the integration of the local orientations based on surface representation yields the curvature selectivity.

It is expected that the activities of model V4 neurons that showed the curvature selectivity are sparse similar to the basis functions. I measured lifetime sparseness of the model V4 neurons and that of the basis functions. The distributions of lifetime sparseness were identical, suggesting that sparseness and surface representation play crucial role in the integration of responses in V1 and V2, so as to establish the curvature representation in V4.

### 5.1.3 Mechanism of Medial Axis representation

I investigated the integration manner of signals from BO selective neurons for the construction of MA representation. I took into account for the onset synchronization (synchronization caused by stimulus onset) of BO selective neurons. I developed the biologically detailed computational model, and carried out the simulations with stimuli including natural objects. The model reproduced the characteristics of the V1 neurons in terms of the distribution of activities and their latency. In addition, simulation results showed that the model produces the fairly correct MA representation of arbitrary

shapes.

The degree of synchronization seems to be a crucial factor for the formation of MA representation. To confirm this idea, I performed simulations and a psychophysical experiment in which the contours of stimuli were presented somewhat asynchronous manner. Each of the results showed that the degree of synchronization causes the bias in the formation of representation and the human perception of DOF. These results support the proposed hypothesis that the onset synchronization of BO selective neurons is crucial for the establishment of MA representation.

## 5.2 Contributions of the thesis

### 5.2.1 Spatial pooling and sparseness in cortical network

I developed two distinct computational models for the construction of the curvature selectivity, which utilize sparse coding and spatial pooling, respectively. Both of the models reproduced the characteristics of V4 neurons, and showed comparable sparseness. These results suggest that the spatial pooling and sparseness play important roles in the cortical processing, so as to establish the representation of curvature.

Spatial pooling and sparseness have been reported theoretically to be the key to produce shape representation [59, 60]. For example, Lee et al. have proposed the model that is composed of multiple layers [60]. The unit in each layer pools the activities within a small region in the descendent layer (i.e. spatial pooling). They assumed that the representation is regularized in order to generate sparse representation. They learned the optimal representation in each layer that is useful to represent objects. For instance, each unit in the third layer represented a face when face images are fed into

the model. Although the model produces the representation of shape, they did not make the comparison to the physiology. The significance of the thesis is that sparseness and spatial pooling are crucial for the establishment of the selectivity reported by physiological studies in addition to the construction of the basis functions that produce shape representation.

Cadieu et al. have reported that spatial pooling could yield the curvature selectivity (see also section 3.6.2; [31]). Their study was data-oriented, in which the optimal pooling is determined so as to reproduce the physiological data. It is natural to expect that such model reproduce the curvature selectivity. In my biological V4 model, parameters were not chosen for reproducing specific physiological results such as the selectivity for curvature. Given that all combinations, some model cells showed the selectivity, and others did not. It enabled me to investigate that what constraint in the pooling is essential for the construction of the selectivity. The advance in the present study is that the model gives insight into the fundamental neural mechanism involved in cortical networks. Specifically, the model elucidates the crucial role of surface representation on the establishment of the curvature selectivity.

### 5.2.2 A simple mechanism for the integration of BO

The psychophysical experiment showed the bias in the perception of DOF toward a highly synchronized side. This result can be explained as binding of contour components by common fate that is known as one of the Gestalt factors [54-56]. The participants tended to perceive the region as figure when surrounded contour blinked more coherently than the opposite side. Although the underlying neural mechanism of common fate has not been revealed, this study may provide insight into the mechanism

that the synchronized feedback signals from higher cortical regions may be useful for assessing coherency of the contours.

## 5.3 Directions for future work

### 5.3.1 Hierarchical representation

In the present component analysis model, the RFs of model V2 neurons were defined by the combination of two Gabor filters. What will happen if the RFs of model V2 neurons are learned from natural images? It would be great interest to reveal whether the RFs of V2 neurons show the angle selectivity [3, 12, 21]. If the learned basis functions of V2 with appropriate sparseness reproduce the angle selectivity, it suggests that sparse coding is the coding scheme shared in the ventral stream, not limited in the specific cortical areas.

### 5.3.2 Constraints on sparse coding

In the component analysis model, the input is binarized and includes one object within an image, so as to focus solely on shape information. In natural vision, however, the stimuli usually contain rich information in terms of containing color, texture, shading, and multiple objects. Although I restricted myself to use binarized images in the present study, further investigation is needed to clarify that what basis functions (RFs) are learned under the natural viewing.

As described in section 2.6.3, a learning algorithm of sparse code is other crucial factor for the construction of the curvature selectivity. One important direction of future work is to clarify the form of cost function (i.e. the assumption of a probability distribution). The preliminary result showed that the basis functions are not leaned when

error of reconstruction is defined by $l_1$-norm. It implies that the assumption of probability distribution in the cost function may be the important factor for the generation of the curvature selectivity.

### 5.3.3 **Reads out MA representation**

It is still covered that what cortical regions receive MA responses in V1 and establish the neural representation of shape from them. The model of V1-V2 networks including BO selective neurons constructed the MA responses in V1. The MA responses seem to send to higher cortical regions beyond V3 in accordance with the cortical hierarchy [61, 62]. Physiological study has reported that IT neurons encode the shape of an object by the configuration of several MA components together with their surfaces [62]. It is suggested that V1 establishes the local MA representations, and higher cortical region integrates them in order to establish the cortical representation of objects.

# Appendix A. Mathematical description of the biological model

The activity of a single model V4 cell is computed by the spatial summation of V1 cells' activity. The model computes the linear summation of the activities of specific V1 cells ($O_{\theta_i}^{V1}$). I defined the response of single model V4 neuron as:

$$T^{V4} = \sum_{i=1}^{n} \left( G_{\varphi_i} * O_{\theta_i}^{V1} \right)(x, y), \tag{A-1}$$

$$\theta_i \in \{0°, 22.5°, \cdots, 337.5°\}, \tag{A-2}$$

$$\varphi_i \in \{0°, 22.5°, \cdots, 337.5°\}, \tag{A-3}$$

where $G$ represents a Gaussian defined by angular position ($\varphi_i$). The center of $G$ is positioned at approximately 1 degree away from the center of stimulus with angular position $\varphi_i$. Standard deviation ($\sigma$) of $G$ is 10pixels, approximately 0.5 degree in visual angle. $\theta_i$ and $\varphi_i$ are orientation and angular position, respectively. Those are defined by one out from 16 orientations (eq. A-2) and positions (eq. A-3). $n$ represents the number of types of integrated V1 neurons. In the present study, I set $n$ as two. The response passed through a sigmoidal function to realize nonlinearity. I defined the output of the model ($O^{V4}$) as:

$$O^{V4} = sig(T^{V4}), \tag{A-4}$$

$$sig(x) = {a}/{(1 + e^{-(x-s)g})}, \quad \text{(A-5)}$$

where $s$, $g$ and $a$ are constants that determine origin, slope and asymptote of the sigmoidal function, respectively. I set these constants empirically as $s$ = 0.5, $g$ = 10, and $a$ = 1 so as to realize the compressive nonlinearity and limit the output within a range (from 0 to 1).

# Appendix B. Mathematical description of the MA model

## B.1 DOF determination

I adapted Sakai & Nishimura's proposal to model BO selective neurons [49]. In their proposal, a model BO selective neuron has asymmetric facilitative/suppressive regions with respect to its cRF. The activities of model BO neurons were modulated depending on whether stimulus projected onto the facilitative or suppressive regions. I determined the activities of model BO selective neurons at time $t$ as below:

$$O^2(x_1, y_1, t) = input(x_1, y_1) + c \sum_{x,y} \{E(x, y, t-d) + I(x, y, t-d)\}. \quad \text{(B-1)}$$

$(x, y)$ and $(x_1, y_1)$ indicate positions. The first and second terms on the right hand side of the equation are the current for the cRF and its surrounds, respectively. $c$ represents the static weight. $d$ is the synaptic delay that increase as the function of the Euclidean distance between $(x, y)$ and $(x_1, y_1)$. $E$ and $I$ are Excitatory and Inhibitory Post Synaptic Potential (EPSP and IPSP), respectively. I computed EPSP and IPSP by following formulas:

$$E(x, y, t-d) = w(v-e)\{exp(-(t-d)/\tau_{decay}^{exc}) - exp(-(t-d)/\tau_{rise}^{exc})\}, \quad \text{(B-2)}$$

$$I(x, y, t-d) = w(v-e)\{exp(-(t-d)/\tau_{decay}^{inh}) - exp(-(t-d)/\tau_{rise}^{inh})\}. \quad \text{(B-3)}$$

$w$ indicates the weight of the connection determined by a Gaussian function, $v$ and $e$

are the membrane potential and reversal potential, respectively. $\tau_{rise}^{exc}$ ($\tau_{rise}^{inh}$) and $\tau_{decay}^{exc}$ ($\tau_{decay}^{inh}$) denote the time constant for the rise and decay of EPSP (IPSP), respectively. The values of time constants are described in Table 4 [47, 48].

**Table 4. Time constants of EPSP and IPSP.**

| Parameter | Value (ms) |
|---|---|
| $\boldsymbol{\tau_{rise}^{exc}}$ | 0.09 |
| $\boldsymbol{\tau_{decay}^{exc}}$ | 1.5 |
| $\boldsymbol{\tau_{rise}^{inh}}$ | 0.1 |
| $\boldsymbol{\tau_{decay}^{inh}}$ | 50 |

## B.2 Algorithm for reconstruction

Below describes an algorithm for shape reconstruction. The reconstruction was accomplished by the following procedures. I obtained the V1 cells that responded to the MA of a shape:

$$M(x,y) = \begin{cases} 1\ (if\ t \geq\ t_{threshold}) \\ 0\ (otherwise) \end{cases}, \tag{B-4}$$

where $t$ indicates the latency of a V1 cell. $t_{threshold}$ denotes the threshold latency of V1 cells responded to the MA. $t_{threshold}$ is set to 77 ms. Next, I sought the cell that showed the strongest response at each spatial location $(x, y)$. The number of spikes and SD of the integration field were obtained as follows:

$$N(x,y) = \max_{\sigma}\left(S_{0.7}(x,y), S_{2.1}(x,y), S_{3.5}(x,y)\right), \tag{B-5}$$

$$\sigma(x,y) = \begin{cases} \mathrm{argmax}_{\sigma}\left(S_{0.7}(x,y), S_{2.1}(x,y), S_{3.5}(x,y)\right) & (if\ M(x,y) = 1) \\ 0 & (otherwise) \end{cases}. \tag{B-6}$$

$\sigma$ is the size of the integration field of V1 cells at $(x, y)$. Then, I superimposed the Gaussians defined by position of the center $(x_1, y_1)$ and SD ($\sigma(x_1, y_1)$) with weight ($w_{\sigma}$; Table 5):

$$T(x,y) = \sum_{x_1,y_1} N(x_1,y_1) \times g_{x_1,y_1}(x,y), \tag{B-7}$$

$$g_{x_1,y_1}(x,y) = \frac{w_{\sigma(x_1,y_1)}}{2\pi\sigma(x_1,y_1)} exp\left(-\frac{(x-x_1)^2 + (y-y_1)^2}{\sigma(x_1,y_1)^2}\right). \tag{B-8}$$

Finally, I passed $T$ through the sigmoidal function so as to realize a non-linearity in the reconstruction. The reconstructed image ($R$) is computed as follows:

$$R(x,y) = \frac{1}{1 + exp(-(T(x,y) - threshold)slope)}. \tag{B-9}$$

The $threshold$ and $slope$ are the origin and steepness of the sigmoidal function, respectively. I set $slope = 300$ and $threshold = 0.3 \times MAX$ ($MAX$ represents the

maximum value of the $R$), except for the rounded stone ($threshold = 0.4 \times MAX$), and the L-shaped tree branch and U-shaped object ($threshold = 0.5 \times MAX$).

**Table 5. Weights for reconstruction. Subscripts of w indicate the spatial extent (SD) of integration fields.**

| SD | Weight |
|---|---|
| $w_{0.7}$ | 0.6 |
| $w_{2.1}$ | 1.0 |
| $w_{3.5}$ | 1.5 |

# Bibliography


1. D. J. Felleman and D. C. Van Essen, "Distributed Hierarchical Processing in the Primate Cerebral Cortex," *Cerebral Cortex,* vol. 1, pp. 1-47, 1991.
2. D. H. Hubel and T. N. Wiesel, “Receptive fields and functional architecture of monkey striate cortex”, *Journal of Physiology*,195, 215-243, 1968
3. M. Ito and H. Komatsu, "Representation of angles embedded within contour stimuli in area V2 of macaque monkeys," *Journal of Neuroscience,* vol. 24, pp. 3313-3324, 2004.
4. H. Zhou, H. S. Friedman, and R. von der Heydt, "Coding of border ownership in monkey visual cortex," *Journal of Neuroscience,* vol. 20, pp. 6594-6611, 2000.
5. A. Pasupathy and C. E. Connor, "Responses to contour features in macaque area V4," *Journal of Neurophysiology,* vol. 82, pp. 2490-2502, 1999.
6. A. Pasupathy and C. E. Connor, "Shape representation in area V4: Position-specific tuning for boundary conformation," *Journal of Neurophysiology,* vol. 86, pp. 2505-2519, 2001.
7. A. Pasupathy and C. E. Connor, "Population coding of shape in area V4," *Nature Neuroscience,* vol. 5, pp. 1332-1338, 2002.
8. E. T. Carlson, R. J. Rasquinha, K. C. Zhang, and C. E. Connor, "A Sparse Object

Coding Scheme in Area V4," *Current Biology,* vol. 21, pp. 288-293, 2011.

9. Y. Yamane, E. T. Carlson, K. C. Bowman, Z. H. Wang, and C. E. Connor, "A neural code for three-dimensional object shape in macaque inferotemporal cortex," *Nature Neuroscience*, vol. 11(11), pp. 1352-1360, 2008
10. T. S. Lee, D. Mumford, R. Romero, and V. A. F. Lamme, "The role of the primary visual cortex in higher level vision," *Vision Research,* vol. 38, pp. 2429-2454, 1998.
11. S. H. Kim and J. Feldman, "Globally inconsistent figure/ground relations induced by a negative part," *Journal of Vision,* vol. 9, 2009.
12. M. Ito and N. Goda, "Mechanisms underlying the representation of angles embedded within contour stimuli in area V2 of macaque monkeys," *European Journal of Neuroscience,* vol. 33, pp. 130-142, 2011.
13. B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature,* vol. 381, pp. 607-609, 1996.
14. J. P. Jones and L. A. Palmer, "The Two-dimensional Spatial Structure of Simple Receptive-fields in Cat Striate Cortex," *Journal of Neurophysiology,* vol. 58, pp. 1187-1211, 1987.
15. N. C. Rust and J. J. DiCarlo, "Balanced Increases in Selectivity and Tolerance Produce Constant Sparseness along the Ventral Visual Stream," *Journal of Neuroscience,* vol. 32, pp. 10170-10182, 2012.
16. B. D. B. Willmore, R. J. Prenger, and J. L. Gallant, "Neural Representation of Natural Images in Visual Area V2," *Journal of Neuroscience,* vol. 30, pp. 2102-2114, 2010.
17. M. R. DeWeese, M. Wehr, and A. M. Zador, "Binary spiking in auditory cortex,"

*Journal of Neuroscience,* vol. 23, pp. 7940-7949, 2003.

18. D. M. Schneider and S. M. N. Woolley, "Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes," *Neuron,* vol. 79, pp. 141-152, 2013.
19. R. A. Jortner, S. S. Farivar, and G. Laurent, "A simple connectivity scheme for sparse coding in an olfactory system," *Journal of Neuroscience,* vol. 27, pp. 1659-1669, 2007.
20. R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, and I. Fried, "Invariant visual representation by single neurons in the human brain," *Nature,* vol. 435, pp. 1102-1107, 2005.
21. A. Anzai, X. M. Peng, and D. C. Van Essen, "Neurons inmonkey visual area V2 encode combinations of orientations," *Nature Neuroscience,* vol. 10, pp. 1313-1321, 2007.
22. J. M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders, "The Amsterdam Library of Object Images," *International Journal of Computer Vision,* vol. 61, pp. 103-112, 2005.
23. C. M. Bishop, “Pattern Recognition and Machine Learning,” *Springer*, 2006
24. B. M. Dow, A. Z. Snyder, R. G. Vautin, and R. Bauer, "Magnification Factor and Receptive-Field Size in Foveal Striate Cortex of the Monkey," *Experimental Brain Research,* vol. 44, pp. 213-228, 1981.
25. R. Gattass, C. G. Gross, and J. H. Sandell, "Visual Topography of V2 in the Macaque," *Journal of Comparative Neurology,* vol. 201, pp. 519-539, 1981.
26. R. Gattass, A. P. B. Sousa, and C. G. Gross, "Visuotopic Organization and Extent of V3 and V4 of the Macaque," *Journal of Neuroscience,* vol. 8, pp. 1831-1845, Jun 1988.

27. D. H. Hubel, and T. N. Wiesel, “Receptive Fields, Binocular Interaction and Functional Architecture in the Cat’s Visual Cortex,” *Journal of Physiology*, vol.160, pp. 106-154, 1962
28. H. R. Wilson and F. Wilkinson, "Detection of global structure in Glass patterns: implications for form vision," *Vision Research,* vol. 38, pp. 2933-2947, 1998.
29. T. Serre, A. Oliva, and T. Poggio, "A feedforward architecture accounts for rapid categorization," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 104, pp. 6424-6429, 2007.
30. A. J. Rodriguez-Sanchez and J. K. Tsotsos, "The Roles of Endstopped and Curvature Tuned Computations in a Hierarchical Representation of 2D Shape," *Plos One,* vol. 7, 2012.
31. C. Cadieu, M. Kouh, A. Pasupathy, C. E. Connor, M. Risenhuber, and T. Poggio, “A Model of V4 Shape Selectivity and Invariance,” *Journal of Neurophysiology*, Vol. 98, pp.1733-1750, 2007
32. M. Carandini, D. J. Heeger, and J. A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *Journal of Neuroscience,* vol. 17, pp. 8621-8644, Nov 1997.
33. W. E. Vinje and J. L. Gallant, "Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1," *Journal of Neuroscience,* vol. 22, pp. 2904-2915, 2002.
34. H. Blum, “Biological Shape and Visual Science (part I),” *Journal of Theoretical Biology*, Vol. 38, pp. 205-287, 1973.
35. D. Marr, and H. K. Nishihara, “Representation and recognition of the spatial organization of three-dimensional shapes,” *Proceedings of the Royal Society of*

*London. Series B Biological Science*, Vol. 200, pp. 269-294, 1978.

36. B. B. Kimia, “On the role of medial geometry in human vision,” *Journal of Physiology Paris*, Vol. 97, pp. 155-190, 2003.
37. I. Kovacs, A. Feher, and B. Julesz, "Medial-point description of shape: a representation for action coding and its psychophysical correlates," *Vision Research,* vol. 38, pp. 2323-2333, 1998.
38. B. B. Kimia, "On the role of medial geometry in human vision," *Journal of Physiology Paris,* vol. 97, pp. 155-190, 2003.
39. Y. Dong, S. Mihalas, F. T. Qiu, R. von der Heydt, and E. Niebur, "Synchrony and the binding problem in macaque visual cortex," *Journal of Vision,* vol. 8, 2008.
40. M. A. Smith and A. Kohn, "Spatial and Temporal Scales of Neuronal Correlation in Primary Visual Cortex," *Journal of Neuroscience,* vol. 28, pp. 12591-12603, 2008.
41. Z. Zhou, M. R. Bernard, and A. B. Bonds, "Deconstruction of spatial integrity in visual stimulus detected by modulation of synchronized activity in cat visual cortex," *Journal of Neuroscience*, vol. 28, pp. 3759-3768, 2008.
42. V. Bringuier, F. Chavane, L. Glaeser, and Y. Fregnac, "Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons," *Science,* vol. 283, pp. 695-699, 1999.
43. A. Angelucci, J. B. Levitt, E. J. S. Walton, J. M. Hupe, J. Bullier, and J. S. Lund, "Circuits for local and global signal integration in primary visual cortex," *Journal of Neuroscience,* vol. 22, pp. 8633-8646, 2002.
44. P. Girard, J. M. Hupe, and J. Bullier, "Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities," *Journal of Neurophysiology,* vol. 85, pp. 1328-1331, 2001.

45. M. L. Hines and N. T. Carnevale, "The NEURON simulation environment," *Neural Computation,* vol. 9, pp. 1179-1209, 1997.
46. A. L. Hodgkin, and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *Journal of Physiology*, Vol. 117, pp. 500-544, 1952.
47. W. Gerstner, and W. Kistler, “Spiking Neuron Models: Single Neurons, Populations, Plasticity,” Cambridge University press, 2002
48. K. A. Archie, and B. W. Mel, “A model for intradendritic computation of binocular disparity,” *Nature Neuroscience*, 3(1), pp. 54-63, 2000.
49. K. Sakai and H. Nishimura, "Surrounding suppression and facilitation in the determination of border ownership," *Journal of Cognitive Neuroscience,* vol. 18, pp. 562-579, 2006.
50. H. E. Jones, W. Wang, and A. M. Sillito, "Spatial organization and magnitude of orientation contrast interactions in primate V1," *Journal of Neurophysiology,* vol. 88, pp. 2796-2808, Nov 2002.
51. E. T. Rolls and G. Deco, “Computational neuroscience of Vision,” *Oxford University Press*, 2002.
52. X. Huang and M. A. Paradiso, "V1 response timing and surface filling-in," *Journal of Neurophysiology,* vol. 100, pp. 539-547, Jul 2008.
53. C. C. Fowlkes, D. R. Martin, and J. Malik, "Local figure-ground cues are valid for natural images," *Journal of Vision,* vol. 7, 2007.
54. P. J. B. Hancock, L. Walton, G. Mitchell, Y. Plenderleith, and W. A. Phillips, "Segregation by onset asynchrony," *Journal of Vision,* vol. 8, 2008.
55. A. B. Sekuler and P. J. Bennett, "Generalized common fate: Grouping by common luminance changes," *Psychological Science,* vol. 12, pp. 437-444, 2001.

56. M. Usher and N. Donnelly, "Visual synchrony affects binding and segmentation in perception," *Nature,* vol. 394, pp. 179-182, 1998.
57. L. Zhaoping, “V1 mechanisms and some figure–ground and border effects,” *Journal of Physiology Paris*, Vol. 97, pp. 503-515, 2003.
58. A. Cowey, E. T. Rolls, “Human cortical magnification factor and its relation to visual acuity,” *Experimental Brain Research*, Vol. 21, pp. 447-454, 1974.
59. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances in Neural Information Processing Systems 25*, pp.1106-1114, 2012
60. H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, “Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations,” *Proceedings of the 26$^{th}$ Annual International Conference on Machine Learning*, pp.609-616, 2009
61. M. D. Lescroart and I. Biederman, "Cortical Representation of Medial Axis Structure," *Cerebral Cortex,* vol. 23, pp. 629-637, 2013.
62. C. C. Hung, E. T. Carlson, and C. E. Connor, "Medial Axis Shape Coding in Macaque Inferotemporal Cortex," *Neuron,* vol. 74, pp. 1099-1113, 2012.

# List of publications

### *Refereed Journals*

1. Y. Hatori and K. Sakai, “Early Representation of Shape by Onset Synchronization of Border-Ownership-Selective Cells in the V1–V2 Network,” *Journal of the Optical Society of America A*, accepted

### *Refereed Proceedings*

2. Y. Hatori, T. Mashita, and K. Sakai, “Sparseness Controls the Receptive Field Characteristics of V4 Neurons: Generation of Curvature Selectivity in V4,” *Lecture Notes in Computer Science*, Vol. 8131, pp.327-334, 2013
3. Y. Hatori and K. Sakai, “Surface-based Construction of Curvature Selectivity from the Integration of Local Orientations,” *Lecture Notes in Computer Science*, Vol.7665, pp.425-433, 2012
4. Y. Hatori and K. Sakai, “Robust Detection of Medial-Axis by Onset Synchronization of Border-Ownership Selective Cells and Shape Reconstruction from its Medial-Axis,” *Lecture Notes in Computer Science*, Vol.5506, pp.301-308, 2009
5. Y. Hatori and K. Sakai, “Representation of Medial Axis from Synchronous Firing of Border-Ownership Selective Cells,” *Lecture Notes in Computer Science*, Vol.4984,

pp.18-26, 2008

### *Proceedings of Conference*

6. Y. Hatori and K. Sakai, “Surface Representation in Early Visual Cortices for the Construction of Curvature Selectivity in V4,” *IEICE Technical Report*, Vol.112, No.480, pp.73-78, 2013
7. Y. Hatori and K. Sakai, “Medial-Axis Representation of Shape from Onset Synchronization,” *IEICE Technical Report*, Vol.108, No.480, pp.153-157, 2009
8. Y. Hatori and K. Sakai, “A medial axis representation by neural population for shape perception,” *Technical Report of the Institute of Image Information and Television Engineering*, Vol.31, No.19, 2007, pp29-32

### *Conference Abstracts*

9. Y. Hatori, T. Mashita, and K. Sakai, “Sparse Representation in the Construction of Curvature Selectivity in V4”, *36th European Conference on Visual Perception 2013*, P16-177
10. Y. Hatori, T. Mashita, and K. Sakai, “Surface Based Integration of Local Orientations for the Construction of Curvature Selectivity -Neural Mechanism of Sparse Coding in V4-,” *Neuro2013*, P2-2-100
11. Y. Hatori and K. Sakai, “Curvature selectivity Based on the Surface Representation in Lower- to Intermediate-level Visual Cortices”, *Society for Neuroscience Annual meeting 2012*
12. Y. Hatori and K. Sakai, “Modeling Curvature Synthesis Based on the Surface Representation in Early to Intermediate Visual Areas,” *The 35th Annual Meeting of*

*the Japanese Neuroscience Society 2012*, P2-a36

13. S. Matsuoka, Y. Hatori, and K. Sakai, “Perception of Border Ownership by Multiple Gestalt Factors”, *The 8th Asia-pacific Conference on Vision 2012*, O4-6
14. Y. Hatori and K. Sakai, “Roles of Surface Representation in Early Visual Areas for the Construction of Curvature Selectivity”, *Brain Inspired Computing 2012*, 1-4
15. Y. Hatori and K. Sakai, “A Computational Study on the Representation of Curvature Constructed from Surface-based Integration”, *Vision Science Society Annual Meeting 2012*, *Journal of Vision*, 12(9), 1310
16. Y. Hatori and K. Sakai, “Onset Synchronization of Border-Ownership Selective Cells for the Primitive Representation of Shape in V1-V2 Recurrent Network,” *The 34th Annual Meeting of the Japanese Neuroscience Society 2011*, P2-i05
17. Y. Hatori and K. Sakai, “Surface Construction from the Onset Synchronization of Border-Ownsership Cells in V1-V2 Model”, *Vision Science Society Annual Meeting 2011*, *Journal of Vision*, 11(11), 886
18. Y. Hatori and K. Sakai, “Robust Representation of Medial Axis by the Onset Synchronization of Border-ownership Selective Cells,” *The 31th Annual Meeting of the Japanese Neuroscience Society 2008*, *Neuroscience Research*, Vol.61, Supplement1, S175