Modeling Neural Networks for Artificial Vision

1 author:

Kunihiko Fukushima
Fuzzy Logic Systems Institute
154 PUBLICATIONS **12,311** CITATIONS

REVIEW

Modeling Neural Networks for Artificial Vision

Kunihiko Fukushima
Graduate School of Informatics, Kansai University
Takatsuki, Osaka 569–1095, Japan
E-mail: fukushima@m.ieice.org

(Submitted on November 23, 2005; Accepted on March 20, 2006)

Abstract — Modeling neural networks is a powerful approach to uncover the mechanism of the brain, and the results of the research are ready to use for engineering applications. This paper introduces several models for vision from recent works by the author.

(1) Extraction of optic flow: In the MST area of the monkey, there are cells that respond selectively to specific motions of a large area of the visual field, such as rotation or expansion/contraction. They respond steadily even when the location of the center of optic flow shifts on the retina. They are thought to analyze optic flows of the retinal images. A neural network model for these cells is designed based on vector field hypothesis. It is a hierarchical multilayered network: retina, layer V1, layers MT and layer MST.

(2) Use of blur for robust image processing: In the process of comparing images, the blurring operation increases robustness against deformations and various kinds of noise, and largely reduces computational cost. As an example of usefulness of blurring operation, the paper discusses a neural network model that extracts axes of symmetry from visual patterns. The model is a hierarchical multi-layered network. The model checks conditions of symmetry, not directly from the oriented edges, but from a blurred version of the response of edge-extracting layer. The input patterns can be complicated line drawings, plane figures or gray-scaled natural images taken by CCD cameras.

(3) A model capable of recognizing and restoring partly occluded patterns: Even the identical image is perceived differently by human beings depending on the shape of occluding objects. The model responds in a similar way as human beings. It is a multi-layered hierarchical neural network, in which visual information is processed by interaction of bottom-up and top-down signals. Occluded parts of a pattern are restored mainly by feedback signals from the highest stage of the network, while the unoccluded parts are reproduced mainly by signals from lower stages. The model does not use a simple template matching method. It can recognize and restore even deformed versions of learned patterns.

Keywords — Vision, Neural network model, Visual motion, Optic flow, Use of blur, Symmetry axis, Occluded pattern, Pattern recognition, Pattern restoration

1. Introduction

Even the most advanced electronic computers today exhibit far less ability than human beings when they are required to do intelligent works. In order to design an intelligent information processing system that has higher functions like the human brain, it is important to understand the mechanism of the brain and learn from the real biological basis of it.

The biological brain, however, has a complicated network architecture and is difficult to understand by only conventional neurophysiological and psychological experiments. Hence, modeling neural networks plays an important role. In the modeling approach, we study how to interconnect neurons to synthesize a brain model, or a network with the same functions and abilities as the brain. When synthesizing a model, we try to follow physiological evidence as faithfully as possible. For parts that are not yet clear, we construct a hypothesis and synthesize

a model that follows the hypothesis. We then analyze or simulate the behavior of the model, and compare it with that of the brain. If we find any discrepancy in behavior between the model and the brain, we change the initial hypothesis and modify the model. We then test the behavior of the model again. We repeat this procedure until the model behaves in the same way as the brain. Although we must still verify the validity of the model by physiological experiment, it is probable that the brain uses the same mechanism as the model, because both respond in the same way. The relationship between modeling neural networks and neurophysiology resembles that between theoretical physics and experimental physics.

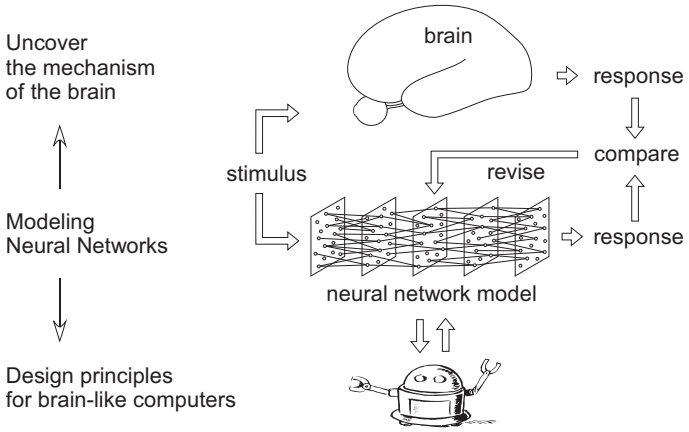


Figure 1. Modeling neural networks

Once we complete a model, we can easily see the essential algorithms of information-processing in the brain, because nonessential factors in the biological brain have been removed from the model. The algorithms of the models are described exactly by mathematical equations or computer programs, which have been necessary to analyze and simulate the behavior of the models. Hence we can use the algorithms, which are ready to run on computers, directly as design principles for intelligent machines of the next generation.

This paper introduces some neural network models from recent works of the author: (1) extraction of optic flow, (2) use of blur for robust image processing, and (3) recognition and restoration of partly occluded patterns.

2. Extraction of Optic Flow

In the visual systems of mammals, information concerning visual motion is mainly analyzed through the occipito-parietal pathway: retina → area V1 (primary visual cortex) → area MT (middle temporal area) → area MST (medial superior temporal area).

In a dorsal part of MST of the monkey, there are cells that respond selectively to specific motions of a large area of the visual field, such as rotation, expansion/contraction and translation [1][2][3][4][5]. These cells are reported to respond steadily even when the location of the center of optic flow shifts on the retina. They are thought to analyze optic flows of the retinal images.

Historically, there have been two major streams of hypotheses proposed so far to explain neural networks extracting optic flow: direction mosaic hypothesis and vector field hypothesis [6].

The direction mosaic hypothesis supposes that a rotation- or expansion/contraction-selective receptive field is constructed receiving signals from a group of direction selective cells, whose preferred directions are arranged circularly or radially. This mechanism alone cannot explain, however, the location-invariant responses of MST cells. To acquire the location-invariance, the receptive field of an MST cell has to be made up with a number of direction-mosaics from different locations of the visual field [1]. This requires a very complicated operations performed by a single cell, which does not necessarily seem biologically plausible.

The vector field hypothesis assumes that the brain analyzes optic flows using operations called *rot* and *div* in the vector field analysis. It has long been pointed out mathematically that optic flows can be extracted with these operations [7]. Neural network models based on this hypothesis, however, have little been proposed so far.

This section discusses a neural network model for MST cells, which follows the vector field hypothesis, and demonstrates the behavior of the model with computer simulation [8].

2.1 Principles of Extracting Optic Flow

Since the rotation and expansion/contraction of the visual field are extracted with a similar mechanism, we will first discuss the principle of extracting rotation.

We assume the moving velocity $\boldsymbol{v} = (v_x, v_y)$ of the scene can be measured at any location (x, y) in the visual field. Here, we define a value ρ with

$$\rho = \frac{\partial v_y}{\partial x} + \left(-\frac{\partial v_x}{\partial y}\right) \tag{1}$$

Incidentally, this value equals the signed absolute value (or length) of a vector obtained by $\mathrm{rot}\,\boldsymbol{v} = \nabla \times \boldsymbol{v}$, which is an operation called *rotation* (or *curl*) in the vector-field analysis.

If an image covering the whole visual field is rotating around a point with a constant angular velocity, ρ becomes constant everywhere in the visual field and is not affected by the shift in location of the center of the rotation. The value of ρ does not change either, even if the coordinate axes (x and y axes) are rotated or the origin of the coordinates is shifted. The sign of ρ indicates the direction of the rotation (counter-clockwise or clockwise), and its absolute value is proportional to the angular velocity of the rotation.

We can then extract the rotation of the whole visual field, if we sum (or take an average of) the values of ρ for all points within a large area S.

$$\bar{\rho} = \sum_{S} \rho \tag{2}$$

Expansion/contraction of optic flow can also be extracted with a similar mechanism. We define δ, which corresponds to a scalar, $\mathrm{div}\,\boldsymbol{v} = (\nabla \cdot \boldsymbol{v})$, obtained by an operation called *divergence* in the vector-field analysis.

$$\delta = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} \tag{3}$$

If the image is expanding radially from (or contracting to) a point, δ becomes always constant everywhere in the visual field and is not affected by the location of the center of the expansion/contraction. A positive value of δ indicates expansion, and a negative value indicates contraction. We then can extract the expansion/contraction of the optic flow, $\bar{\delta}$, by summing (or taking an average of) δ within a large area in the visual field.

2.2 Outline of the Model

The model is a hierarchical multilayered network as shown in Figure 2. It consists of layers of cells connected in a cascade: retina → V1 → $\mathrm{MT_{abs}}$ → $\mathrm{MT_{rel}}$ → MST. Layer V1 corresponds to the primary visual cortex, layers $\mathrm{MT_{abs}}$ and $\mathrm{MT_{rel}}$ to area MT, and layer MST to the dorsal part of area MST of the animal.

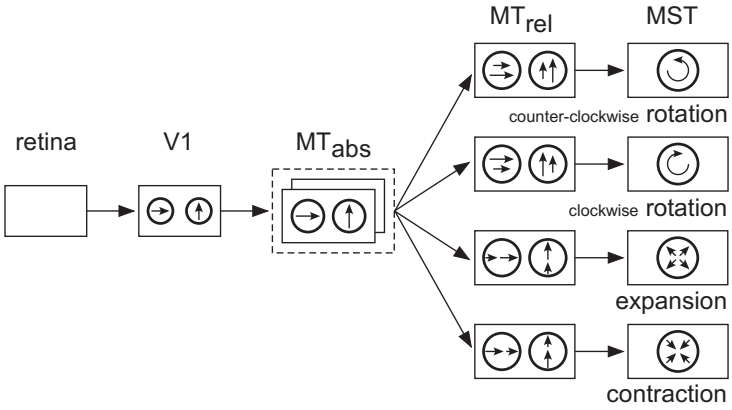


Figure 2. Outline of the architecture of the model for extracting optic flow.

The retina of our model is a photoreceptor array that receives visual image. Each cell of layer V1 has a small receptive field and has direction selectivity. It responds strongly to an object moving in a specific direction. The strength of the response is assumed to be proportional to the velocity of the object.

Since our model uses cells whose responses are proportional to the firing frequencies of biological neurons, their responses take only positive analog values or zero, and cannot be negative. We need a pair of cells to represent the velocities of opposite directions: that is, one for responding to the positive, and the other to the negative velocity. The same is true for other cells in higher stages in the network. We will then use suffix $+$ and $-$ to distinguish cells extracting positive and negative components of ρ, δ, $\bar{\rho}$ and $\bar{\delta}$. Namely, we will use notation such as ρ_+, for example.

2.3 MT-cells

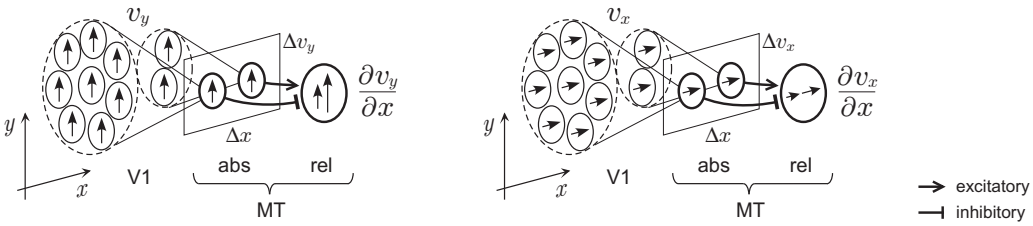
There are two layers for MT-cells, namely, MT_{abs}-cells that extract absolute velocity, and MT_{rel}-cells that extract relative velocity (or velocity gradient).

Cells of layer MT_{abs} respond to the moving velocities of the stimuli in a similar way as V1-cells, but have somewhat larger receptive fields. Each MT_{abs}-cell has excitatory input connections from a group of V1-cells of the same preferred direction but of slightly scattered receptive-field locations.[1]

Cells of layer MT_{rel} work to extract gradients of velocity, which correspond to $\partial v_y/\partial x$, $(-\partial v_x/\partial y)$, $\partial v_x/\partial x$ and $\partial v_y/\partial y$ in Eqs. (1) and (3).

Each MT_{rel}-cell receives antagonistic inputs from two MT_{abs}-cells of the same preferred direction, whose receptive fields adjoins each other. An MT_{rel}-cell thus extracts the gradient of local velocity at a certain location in the visual field.

Figure 3(a) illustrate the network that extracts $\partial v_y/\partial x$, which is used for extracting rotation. It calculates the difference in vertical velocity Δv_y between two points separated horizontally by Δx, by receiving antagonistic inputs from two MT_{abs}-cells of vertical preferred direction, whose receptive field centers differ horizontally.



(a) Network for extracting rotation (b) Network for extracting expansion

Figure 3. A neural network for extracting relative velocity.

If we thus calculate the gradients of velocities in both vertical and horizontal directions and sum them up, we can obtain rotation at an arbitrary point in the visual field. Figure 4, for example, illustrates MT_{rel}-cells relevant to extracting counter-clockwise rotation ρ_+ at an arbitrary location in the visual field. Since MT_{abs}-cells, like V1-cells, cannot generate negative outputs, a pair of cells is used to represent positive and negative components of the velocity in each direction. Gradients of velocity have to be extracted from these four components. The outputs of cells $v_y \geq 0$ and $v_y \leq 0$ in the figure correspond to values of $\partial v_y/\partial x$ extracted from the positive and negative components of the vertical velocity v_y, respectively, and outputs of $v_x \geq 0$ and $v_x \leq 0$ to the value of $(-\partial v_x/\partial y)$ extracted from those of the horizontal velocity.

Each term of Eq. (3), which is used for extracting expansion/contraction, can be obtained by a network like Figure 3(b), which resembles Figure 3(a). The only difference between the two networks resides in the relative locations of the receptive fields of the excitatory and inhibitory MT_{abs}-cells. For the extraction of rotation, they are arranged in the direction orthogonal to their preferred direction, while they are arranged in the direction of their preferred direction for the extraction of expansion/contraction.

In other words, four types of MT_{rel}-cells, which extract rotation in both directions, expansion and contraction, can be constructed by simply changing the location of the inhibitory area relative to the location of the excitatory area as illustrated in Figure 5.

[1] According to recent neurophysiological findings, signals from V1-cells are not the only inputs to direction selective MT-cells. We simply assumed in our model, however, that MT_{abs}-cells receive input connections only from V1-cells, because the main point of discussion of our model is on the network between MT and MST areas.

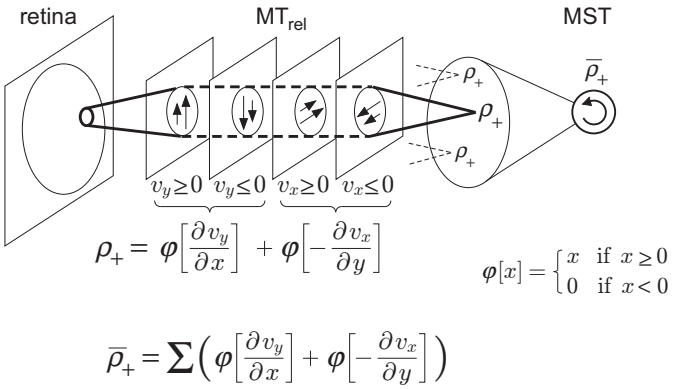


Figure 4. Extracting rotation at a location from four components of relative velocities.

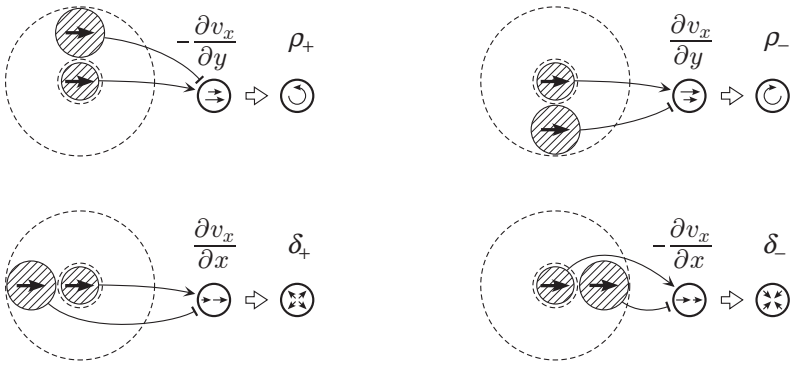


Figure 5: Differences in relative locations between the excitatory and inhibitory areas of four types of MT_{rel}-cells, which extract rotation in both directions, expansion and contraction.

2.4 MST-Cells

Each cells of layer MST gathers outputs of MT_{rel}-cells from a very large area and detect the types of optic flow, such as rotation or expansion/contraction. The optic flow of a large visual field, $\bar{\rho}$ or $\bar{\delta}$, can thus be extracted independently of the location of its center. We write $\bar{\rho}_+$ and $\bar{\rho}_-$ to represent MST-cells extracting counterclockwise ($\bar{\rho} > 0$) and clockwise ($\bar{\rho} < 0$) rotations, respectively. We also write $\bar{\delta}_+$ and $\bar{\delta}_-$ to represent MST-cells extracting expansion ($\bar{\delta} > 0$) and contraction ($\bar{\delta} < 0$).

Mathematically, the operation performed by a counterclockwise-rotation selective MST-cell, for example, can be expressed roughly by

$$\bar{\rho}_+ = \sum_S \left(\varphi\left[\frac{\partial v_y}{\partial x}\right] + \varphi\left[-\frac{\partial v_x}{\partial y}\right] \right) \tag{4}$$

because the outputs of MT_{rel}-cells do not take negative values. Here, $\varphi[\,]$ represents a function that takes the positive component of its argument, namely, $\varphi[x] = \max(x, 0)$.

It should be noted here that an MST-cell adds responses of MT_{rel}-cells directly without introducing interneurons. Although we showed an illustration like Figure 4 above, this does not mean that an independent cells like ρ_+ really exist in the network.

Figure 6 illustrate only the portion of our network that is relevant to extracting counter-clockwise rotation. In the explanations up to here, we have made discussions assuming that the direction selectivity (or the tuning curve) of V1- and MT_{abs}-cells takes a cosine curve. If it follows a cosine curve, the velocity in any direction can be represented by two pairs of cells representing the velocity in two orthogonal directions, namely, by four cells of preferred directions of $0°$, $90°$, $180°$ and $270°$. The direction selectivity of V1- and MT-cells of animals, however, is usually slightly sharper than a cosine curve. We then prepare cells of eight different preferred directions in our model. In other words, our model uses, not only horizontal and vertical components of velocities, but also velocity-components in another set of orthogonal axes rotated by $45°$.

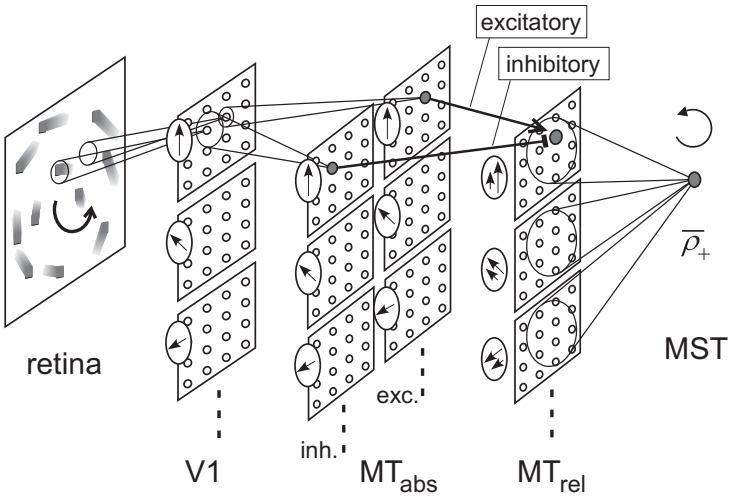


Figure 6. The neural network relevant to extraction of counter-clockwise rotation.

2.5 Computer Simulation

The neural network is simulated on a computer. The stimuli given to the retina are black-and-white random-dot patterns that are moving.[2]

Figure 7 shows a response of the network to a random-dot pattern rotating counter-clockwise. Each square in the figure represent a cell-plane.[3] The output of each cell is shown by the darkness of a dot in the figure.

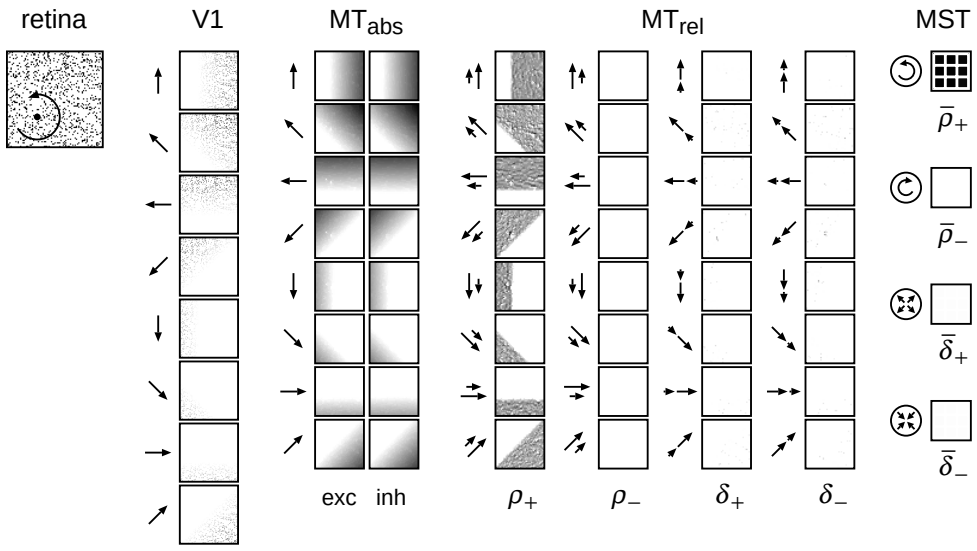


Figure 7. A response of the model to a random-dot pattern rotating counter-clockwise.

Each cell-plane of layer MST consists of nine cells, which have receptive fields of the same characteristics, but at different locations. It can be seen from the figure that all $\bar{\rho}_+$-cells, which extract counter-clockwise rotation, exhibit strong responses independently of the locations of their receptive fields, while all other MST-cells are silent. This means that MST-cells extract the counter-clockwise rotation correctly without being affected by the location of the center of the rotation.

[2]The mechanism of extracting velocity by V1-cells is abbreviated in the present simulation, and the values of the velocities are directly substituted to the outputs of the cells. We assume that the output of each V1-cell is proportional to the velocity of the dot that drops in its receptive field, and that velocity-tuning of V1-cells follows a cosine curve.

[3]A cell-plane is a group of cells that are arranged retinotopically and share the same set of input connections [12]. As a result, they all have receptive fields of an identical characteristic, but the locations of the receptive fields differ from cell to cell.

2.6 Biological Plausibility of the Model

We have shown that rotation and expansion/contraction can be extracted in our model, not by neural networks of different architectures, but by neural networks of the same architecture by simply changing the relative locations between inhibitory and excitatory areas in the receptive fields of MT_{rel}-cells (Figure 5). More specifically, we hypothesize that the difference in the relative location of the inhibitory to the excitatory areas create four different groups of MT_{rel}-cells, which take charge of extracting ρ_+, ρ_-, δ_+ and δ_-; and hence four different types of MST-cells, $\bar{\rho}_+$, $\bar{\rho}_-$, $\bar{\delta}_+$ and $\bar{\delta}_-$.

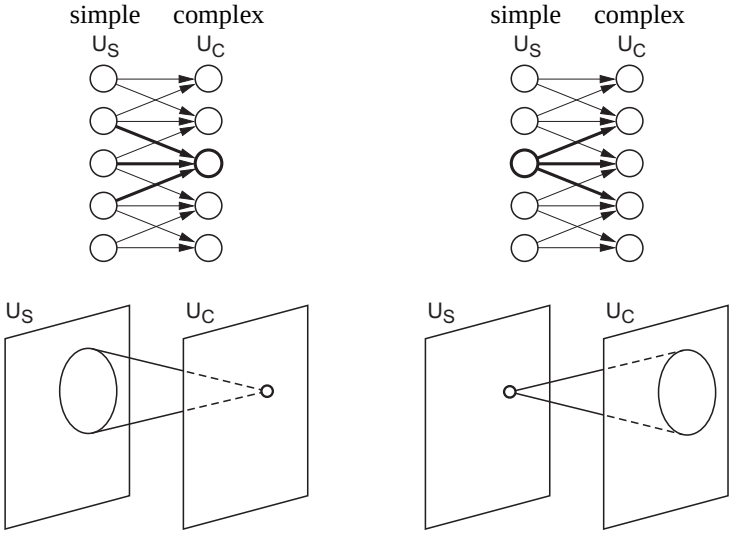
Existence of such MT_{rel}-cells in the biological brain are suggested by several reports on single cell recordings from the visual area MT of macaque monkeys. For example, Tanaka, et al. reported MT-cells that are thought to respond to relative motion [9]. Xiao, et al. reported that, in half of the MT-cells, antagonistic surround was asymmetric, and the inhibitory area was located in a single region on one side of the excitatory area of the receptive field [10].

Graziano, et al. recorded MST neurons that respond preferentially to spiral motions [5]. Our model can also emulate well these MST-cells, if we properly choose MT_{rel}-cells, from which the MST-cell receive signals.

3. Use of Blur for Robust Image Processing: Extraction of Symmetry Axes

It might be felt that a blur decreases the information-processing ability of the visual system. However, this is not always the case. On the contrary, it greatly increases the ability for processing visual information, if the blur is properly controlled.

We mean by the word "blur" such an operation that is performed between simple cells and complex cells in the classical hypothesis by Hubel and Wiesel. A complex cell receives signals from a group of simple cells that have the same receptive field but at slightly scattered locations, as illustrated in Figure 8(a). A response is elicited from the complex cell if at least one of the simple cell is active. Even if a stimulus that elicits a response from a simple cell shift a little, another simple cell will respond instead, and the complex cell will keep responding. This can be interpreted as an operation of tolerating shift in stimulus location. If we look at this operation from other side, it can also be interpreted as a blurring operation. If a simple cell is active, its output is distributed to a group of complex cells as illustrated in Figure 8(b).



(a) Tolerating positional error (b) Blurring responses

Figure 8. Shift-tolerance and blurring operation: Two different interpretation of the same operation.

It is when comparing images that a blur becomes useful. In various kinds of visual information processing, in either biological neural networks or engineering systems, it often becomes necessary to compare visual patterns or images and to judge if they are the same or different. If the goal of the information processing is to recognize visual

patterns, it is necessary to compare an input pattern with a set of learned patterns that have been memorized before. If the goal is to extract an axis of mirror symmetry (bilateral symmetry), we first need to find out a candidate of the axis of symmetry and test whether a pattern is bilaterally symmetrical about the axis, by comparing images located on both sides of the axis.

For these comparisons of images, small amount of distortion of images and noise have to be tolerated. If we want to recognize hand-written characters, an identical pattern would never be written even by the same person. Characters have always some distortions between each other. They will change in scale, in location, in orientation, in brightness, in the width of the lines, and even in shape. If we want to extract symmetry axes from natural images, various noise and distortions would be inevitably contained in input images.

Computational cost is also a big problem in the comparison. Pixelwise comparison, does not only lack tolerance against deformation, but also requires a huge amount of computational cost. It is not plausible that such a computation as pixelwise comparison is performed in the visual system of the biological brain. It is not practical in engineering systems for real world applications, either.

A blurring operation greatly increases robustness against deformations and various kinds of noise, and reduces computational cost. The introduction of blurring operation in neural networks has a long history. For example, the *neocognitron*, which is a neural network model for visual pattern recognition proposed previously by the author [11][12][13][14], contains cells that produce blur and reduce spatial resolution.

This section discusses the usefulness of blur, taking a case of extracting symmetry axes as an example.

3.1 Extraction of Symmetry Axes

Mirror-images, or bilateral symmetries, are perceptually salient, and symmetry about an axis attracts our eyes strongly. Various models or algorithms have been proposed so far for extracting axes of symmetry.

To test whether a pattern is bilaterally symmetrical about an axis, it is necessary to compare, in principle, signals from both side of the axis. Some of the models use pre-wired neural networks or computational methods [15][16], or spatial filters [17] to make this comparison. Some others try to create such neural networks by learning [18][19][20]. Although a variety of methods have been proposed for this comparison, most of them try to use raw signals from input images directly. Direct comparison, however, requires a tremendous amount of computational cost.

Some of the methods try to extract medial axes or skeletons, which are mixtures of axes of global and local symmetries [21][22]. Although they require smaller computational costs, they can process only simple figures, such as plane figures or patterns drawn with single closed curves. The insides of the figures have to be uniform and should not have complicated textures. They have difficulty in processing complicated patterns or natural images.

Introduction of blur is also useful for extracting symmetry axes. We will discuss a neural network that uses blurring operation to extract symmetry axes [23].

3.2 Network Architecture

Our network has a hierarchical multilayered architecture, which resembles that of the lower stages of the neocognitron [12][13][14]. As illustrated in Figure 9, it consists of a number of layers connected in a hierarchical manner: an input layer (U_0), a contrast-extracting layer (U_G), an edge-extracting layer of a simple-type (U_S), an edge-extracting layer of a complex-type (U_C), a symmetry-axis-extracting layer (U_H), and a symmetry-extracting layer (U_X). Each layer consists of a number of cell-planes (See footnote 3 for the explanation of a cell-plane).

The stimulus pattern is presented to input layer U_0, which consists of photoreceptor cells. The output of layer U_0 is fed to contrast-extracting layer U_G, whose cells resemble retinal ganglion cells and have either on-center or off-center receptive fields. The on-center cells extract positive contrast in brightness, whereas the off-center cells extract negative contrast from the image presented to the input layer.

The output of layer U_G is sent to edge-extracting layer (of a simple-type) U_S. Layer U_S consists of S-cells, which resemble simple cells in the primary visual cortex. Each S-cell receives connections from both on- and off-center cells of U_G and extracts edges of a particular orientation [13]. There are K cell-planes in layer U_S, and each cell-plane consists of edge-extracting S-cells of a particular preferred orientation. We will take $K = 32$, in the computer simulation discussed later. Namely, preferred orientations of the cell-planes are determined at an interval of $2\pi/K = 11.25^\circ$. Layer U_S thus decomposes the input image into edge components of various orientations.

The output of layer U_S is fed to layer U_C (edge-extracting layer of complex-type), where a blurred version of

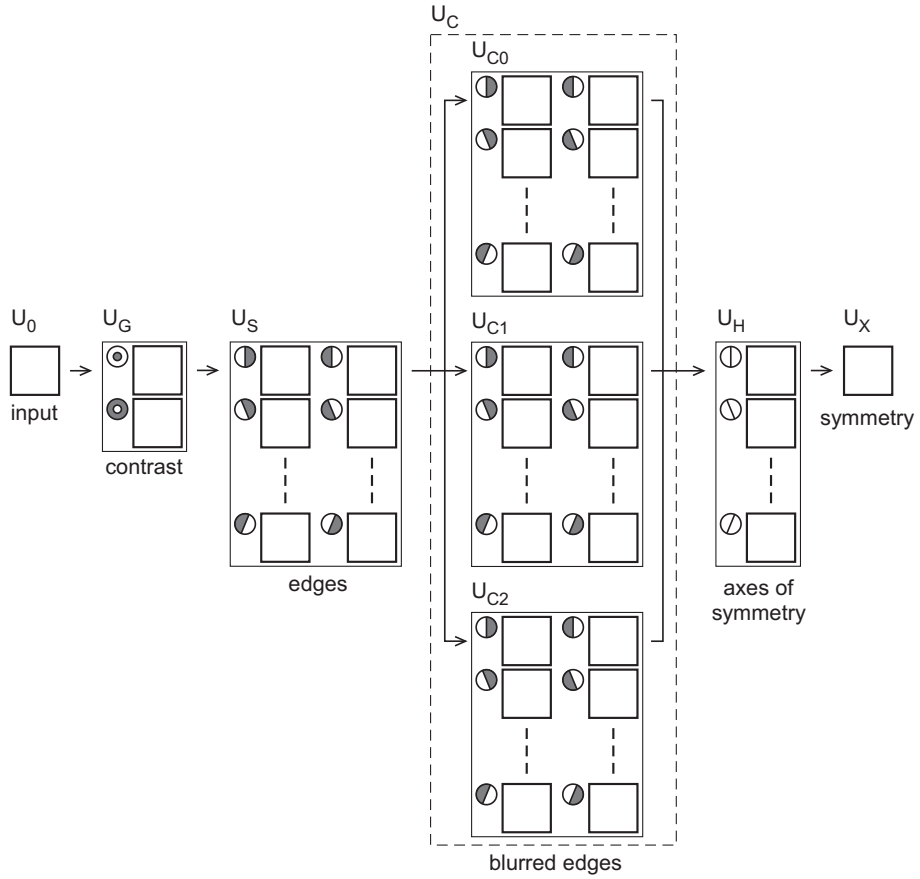


Figure 9. Architecture of the network for extracting symmetry axes.

the response of layer U_S is generated. Layer U_C consists of C-cells, which resemble complex cells in the primary visual cortex. Layer U_C consists of 3 sub-layers, each of which have a different degree of blur.

3.3 Principles of Extracting Symmetry Axis

If a pattern is symmetrical about an axis, a local feature (an oriented edge, in this particular case of our network) on one side of the axis always has its counterpart on the other side of the axis as shown in Figure 10.

We will check if a point, o, is on the axis of symmetry of orientation α. Let an oriented edge exist at a distance A from o in the direction perpendicular to the axis. If the input pattern is symmetrical, another edge should exist at the location symmetrical to the axis, and they make a mirror image to each other. Let the strengths of the edges (namely, the outputs of edge-extracting cells) on the right and left side of the axis be u_r and u_l, respectively. If the input image is symmetrical, u_r and u_l should be equal in strength. We will define a degree of symmetry by $h = \gamma(u_r + u_l) - \delta|u_r - u_l|$. Positive parameters γ and δ determine how much degree of asymmetry be allowed. When δ/γ is small, a small asymmetry can be allowed. When δ/γ is large, h becomes negative if the image is not strictly symmetrical.

If we would measure h for all values of A and for all orientations of edges, and if h be always positive, we could conclude that point o is on an axis of symmetry. This is not practical, however, because of a huge amount of computational cost required.

To reduce the computational cost, we measure h in our network, not directly from the output of edge-extracting layer U_S, but from the output of sub-layers of U_C. Each C-cell (cell of U_C) sums the response of U_S using connections that have a cone-shaped spatial distribution of radius A_C. We then calculate h only at a pair of locations that satisfy $A = A_C$, as is illustrated in Figure 11.

Calculation of h can easily be done by a simple neural circuit, like the one shown in the figure. The circuit

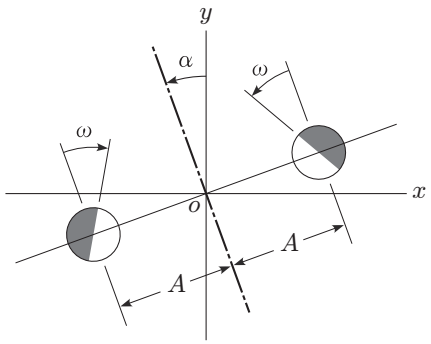


Figure 10. Principles of extracting an axis of symmetry.

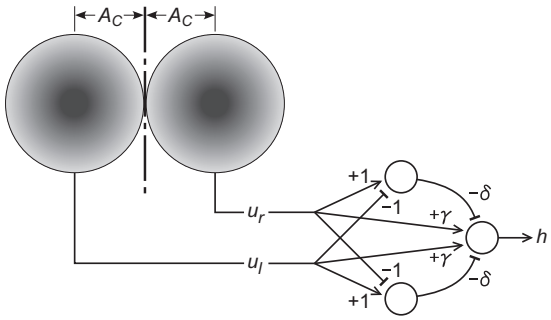


Figure 11: Checking a necessary condition for symmetry from the outputs of C-cells. Lower right of the figure shows an example of a circuit calculating $h = \gamma(u_r + u_l) - \delta|u_r - u_l|$.

consists of analog threshold elements. An analog threshold element takes a weighted sum of input signals, and its output is suppressed if the sum is negative.

Incidentally, the cells at the top and bottom of the circuit come to have receptive fields like hypercomplex cells in the visual cortex, in the sense that each of them receives antagonistic signals from two C-cells (complex cells), whose receptive fields are located adjacently to each other.

3.4 Required Shape of the Blur

It is essential that the shape of input connections to a C-cell, through which the responses of S-cells are gathered, should be non-uniform and of a cone-shape, and should not be flat like a dome. This is illustrated in Figure 12. It is required that the response of a C-cell largely changes by a shift in location of an edge. A signal from a different location of U_S should elicit a different response from the C-cell. In other words, difference in location is encoded into the difference in amplitude of the response of C-cells. If this condition is satisfied, $|u_r - u_l|$ becomes large enough to suppress h, when a pair of edges are not located symmetrically. The values of u_r and u_l themselves are not so important here.

Even when two or more signals come at a time as illustrated in Figure 13, the difference in location of each component of the signals is converted into the difference in amplitude of the response of a C-cell, whose input connections have non-uniform spatial distribution. Asymmetry between the two groups of signals can thus be detected in a single action by the non-uniform blurring operation by the C-cells.

This requirement of non-uniformity of the blur is different from the requirement in the neocognitron for pattern recognition. Shift in location of local features has to be accepted tolerantly for robust recognition of patterns, while it has to be detected sensitively for symmetry axis extraction. In the neocognitron, it is desired that the response of a C-cell does not change largely by the shift of a local feature, and the spatial distribution of the connections to a C-cell is better to be more flat like a dome.

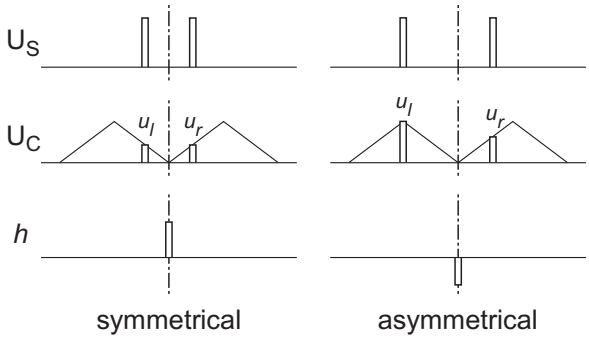


Figure 12: The process of calculating h. When the stimulus is symmetrical, $u_r = u_l$ holds. Hence, h takes a positive value. When the stimulus is asymmetrical, u_r and u_l take different values, and h becomes negative even though the values of u_r and u_l themselves are large, provided δ/γ is large enough.

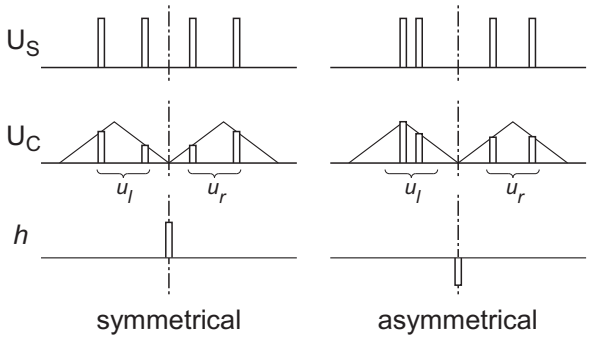


Figure 13: The process of calculating h, when two or more pairs of signals come at a time. Asymmetry between the two groups of signals can be detected in a single action by the non-uniform blurring operation by the C-cells.

3.5 Use of Different Degrees of Blur

As has been discussed above, we need not calculate h for all values of A in Figure 10, if we use the output of U_C instead of U_S. We can simply calculate h only at a pair of locations that satisfy $A = A_C$. To test this condition for all orientations ω of edges, we sum up h for all orientations of edges.

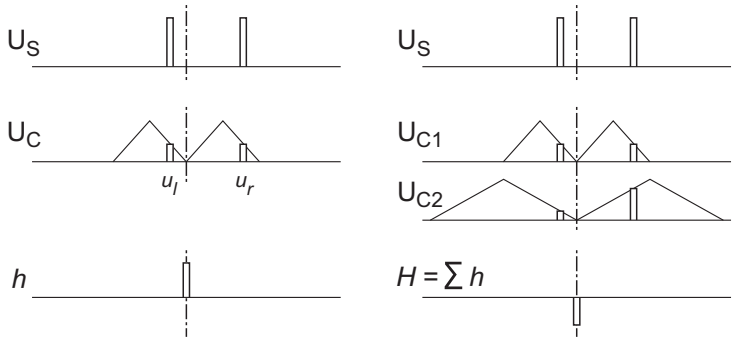
Let the summed h be H. Since each C-cell integrates the output of edge-extracting S-cells within a radius of A_C, we can interpret that H represents the symmetry of edges located within a distance of $2A_C$ from the axis. If there is an asymmetry, H becomes small or negative.

We cannot directly conclude, however, that the input pattern is symmetrical even if H is large, because H represents a *necessary*, but not always a *sufficient*, condition of symmetry. In other words, H always takes a large value on an axis of symmetry, but there is a possibility of taking a large value even at a place that is not on the axis of symmetry in a special occasion. Figure 14(a) shows an example of generation of such a spurious output at a place that is not on the axis of symmetry.

If we use the response of another sub-layer of U_C (say, U_{C2}), however, u_r and u_l take different values, as illustrated in Figure 14(b). In other words, the spurious response is suppressed by combining signals from a number of sub-layers of U_C. Since the summation in the calculation of H is taken without threshold operations for different resolutions, most of the spurious outputs generated only at a single resolution can be suppressed by mutual interaction between M different resolutions.

Therefore, we calculate H with $M\,(=3)$ different values (resolutions) of A_C and sum them up. A circuit for calculating H can easily be made by combining the circuit for calculating h. Since the summation of h is taken without threshold operations, the cell in the right side of Figure 11 are used in common for all calculations of h, and the circuit for calculating H becomes like the one shown in Figure 15.

As a result of M different resolutions and sizes of receptive fields of C-cells, asymmetry near the axis of symmetry is detected more sensitively through the signals from C-cells of higher resolution with smaller receptive fields, while the asymmetry in broader areas is roughly analyzed through the signals from C-cells of larger receptive



(a) A single blur (b) Two different blurs

Figure 14: Use of different degrees of blur to suppress spurious responses. (a) An asymmetrical pattern is erroneously judged as symmetrical, if h is calculated from a single sub-layer of U_C. (b) The spurious response is suppressed by combining signals from two sub-layers of U_C, and the asymmetry can be detected correctly.

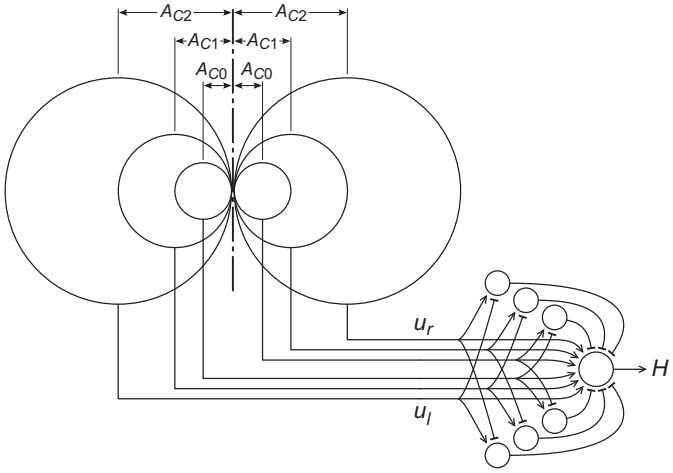


Figure 15: Checking a necessary condition for symmetry from the outputs of C-cells having different degrees of blur. Lower right of the figure shows an example of a circuit calculating $H = \sum h = \sum \{\gamma(u_r + u_l) - \delta|u_r - u_l|\}$.

fields.

This calculation of H is done by cells of layer U_H. Layer U_H consists of $K/2$ cell-planes depending on the orientations of axes of symmetry. Then, symmetry-extracting layer U_X, which consists of only one cell-plane, integrates the responses of all cell-planes of layer U_H. Layer U_X responds at locations of the axes of symmetry of the input image independently of their orientations.

It is important to note here that our network checks conditions of symmetry, not directly from the oriented edges (response of S-cells), but from the response of C-cells, which have larger receptive fields and yield low-resolution responses covering a large area of the input layer. The use of blurred responses from C-cells increases the tolerance for deformation of the input pattern, and greatly reduces computational costs for extracting symmetry axes. It should be noted that the response of a cell-plane of U_C can be used in common for the calculation of h, not only at a single, but at many different orientations of symmetry axes. This helps to reduce the computational cost.

Use of oriented edges (U_C), and not the input image itself (U_0 or U_G), is also useful, when comparing blurred responses. If the input image is directly blurred, most of the important features in the image will be lost. If we use oriented edges, however, most of the important features can still remain even after blurring operations, by which information of edge locations becomes ambiguous.

3.6 Computer simulation

The network is simulated on a computer. Figure 16 shows how the cells in the network respond to an H-shaped pattern.

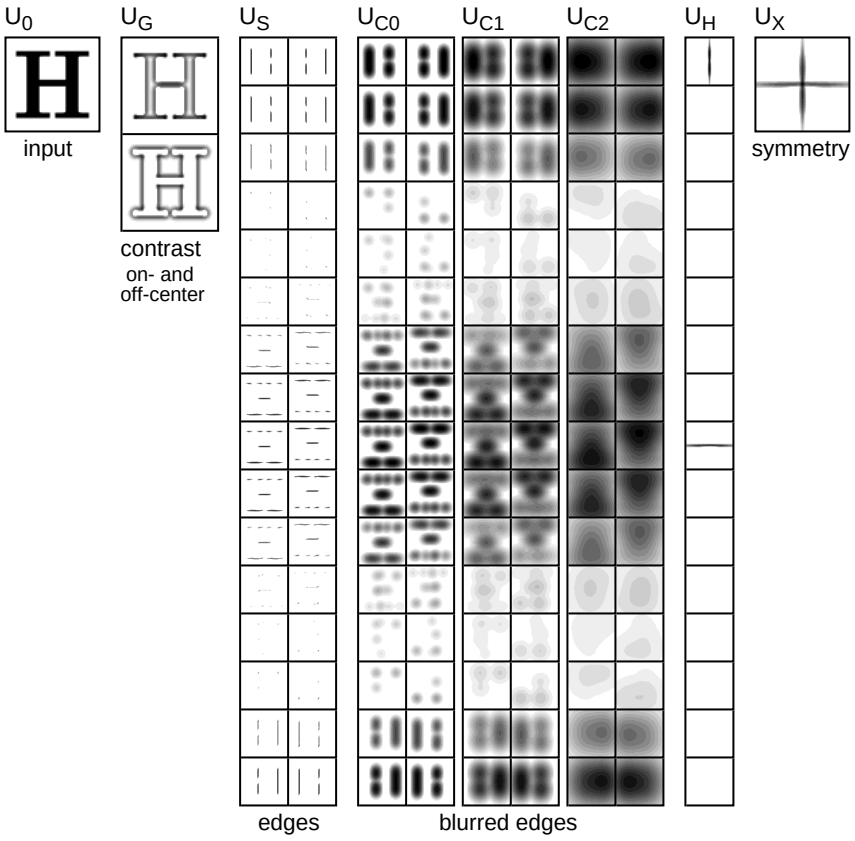


Figure 16. A response of the network that extracts axes of symmetry.

Figure 17 summarize the responses of symmetry-extracting layer U_X for various input patterns given to U_0. As can be seen from these figures, axes of symmetry are extracted correctly even from complicated figures, which have small amount of asymmetry. Our network can extract symmetry axis even from gray-scaled natural images, which are taken by CCD cameras. Even when an input pattern has two or more symmetry axes, they all can be extracted.

Figure 18 shows how robust our network is against various kinds of noise. It can be seen that axes of symmetry are extracted without being affected so much by noises. The input pattern in (a) is an image without noise. The input patterns in (b)–(e) has different levels of additive random noise. The levels of the noise (ratio of r.m.s. noise to peak-to-peak signal) are 0.01, 0.02, 0.05, 0.1, respectively.

The input pattern in (f) has largely reduced brightness contrast in the left of the image. Namely, the contrast at the left side of the car body is about 0.28 of that at the right side. The input pattern in (g) has, not only reduced contrast in the left, but also a shading in the background lighting. The extraction of symmetry axes is not affected so much by these unbalanced lighting conditions. The robustness against shading is obtained partly by the fact that the edges are extracted by layer U_S after extraction of contrast by layer U_G. The robustness against differences in contrast between both sides of the symmetry axis is produced also by the nonlinear (or saturating) input-to-output curve of the C-cells.

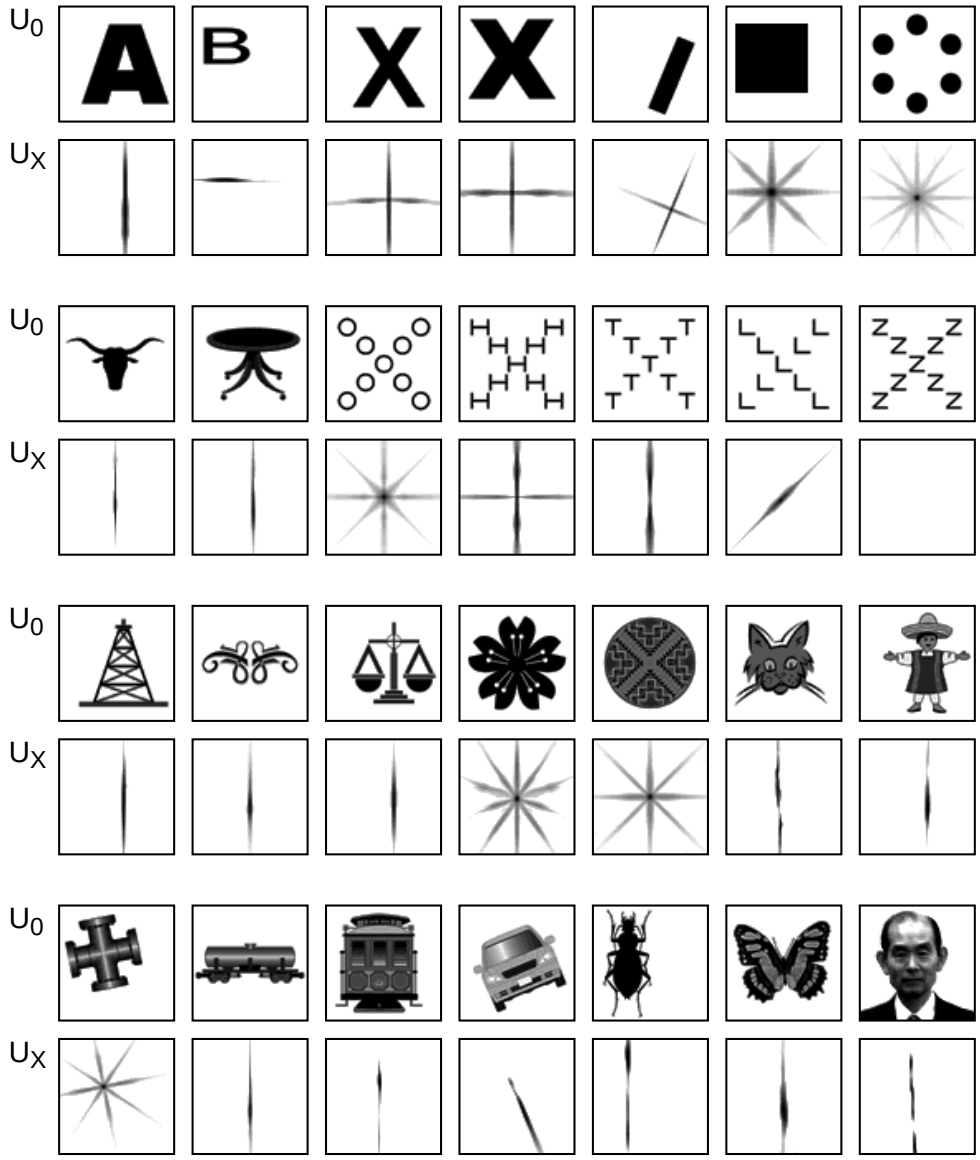


Figure 17. Axes of symmetry extracted from various input patterns.

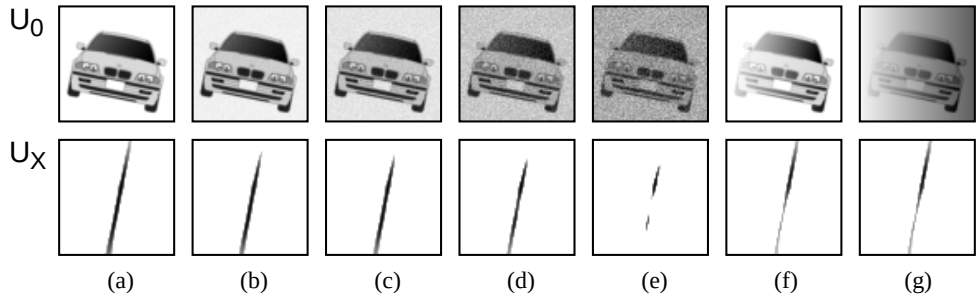


Figure 18: The effect of various kinds of noise on extracting symmetry axes. (a) An original image without noise. (b)–(e) Images with different levels of additive random noise. (f) Reduced contrast in the left. (g) Reduced contrast in the left and a shading in the background.

4. Recognition and Restoration of Partly Occluded Patterns

Human beings are often able to read or recognize a letter or word contaminated by ink stains that partly occlude the letter. If the stains are completely erased and the occluded areas of the letter are changed to white, however, we usually have difficulty in reading the letter, which now has some missing parts. For example, the patterns in Figure 19(a), in which the occluding objects are not visible, are almost illegible, but the patterns in Figure 19(b), in which the occluding objects are visible, are much easier to read.

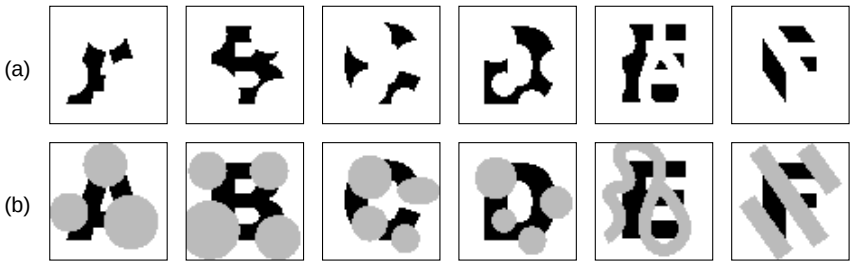


Figure 19. Patterns partly occluded by other objects, which are (a) invisible and (b) visible.

Why does the recognition become easier when the occluding objects are visible? The author assumes the following. Visual patterns have various local features, such as edges and corners. The visual system of animals extracts these features in its lower stages and tries to recognize a pattern using information of extracted local features. When a pattern is partly occluded, a number of new features, which do not exist in the original pattern, are generated near the contour of the occluding objects (See Figure 22). If the occluding objects are not visible, the visual system has difficulty in distinguishing which features are relevant to the original pattern, and which are not. These irrelevant features largely disturb the correct recognition by the visual system.

If the occluding objects are visible, the visual system can easily distinguish relevant from irrelevant features, and can ignore irrelevant features. Since the visual system has a large tolerance to partial absence of relevant features, it can recognize the partly occluded patterns correctly, even though some relevant features are missing.

Figure 20 shows some other examples, in which the perception is largely affected by the placement of occluding objects. In Figure 20(a), the left figure looks as though pattern ‘C’ is partly occluded, but the right figure looks as though pattern ‘O’ is partly occluded. The difference between them resides only in a slight difference in shape of the occluding objects.

In Figure 20(b), the black parts of the patterns are actually identical in shape between the left and right figures. We feel as though different black patterns are occluded by gray objects. Namely, we perceive as though pattern ‘R’ is occluded in the left figure, while pattern ‘B’ is occluded in the right.

(a) (b)

Figure 20. Identical patterns are perceived differently by the placement of different gray objects.

We can assume from these observations, that human beings perceive a break of a line behind an occluding object, not from the fact that a line segment is not visible there, but from the fact that an end (or ends) of a line is (are) visible there. In other words, we perceive a line is connected, if no line-end is visible.

4.1 Neural Network Model

We will introduce a neural network model [24] that shows a response similar to the human-like perception mentioned above.

The model is a multi-layered hierarchical network having forward (bottom-up) and backward (top-down) signal paths. Information is processed in the network through the interaction of forward and backward signals.

The forward paths mainly take charge of recognizing partly occluded patterns, and the backward paths take charge of restoring missing portions of the occluded patterns.

Figure 21 illustrates the network architecture of the model, showing how the information flows between layers. U represents a layer of cells in the forward paths, and W in the backward paths. Each layer consists of cell-planes, in a similar way as in the neocognitron. The architecture resembles that of the *Selective Attention Model* [25], but the feedback signals come down to lower stages, not only from the highest stage, but also from every stages of the hierarchy.

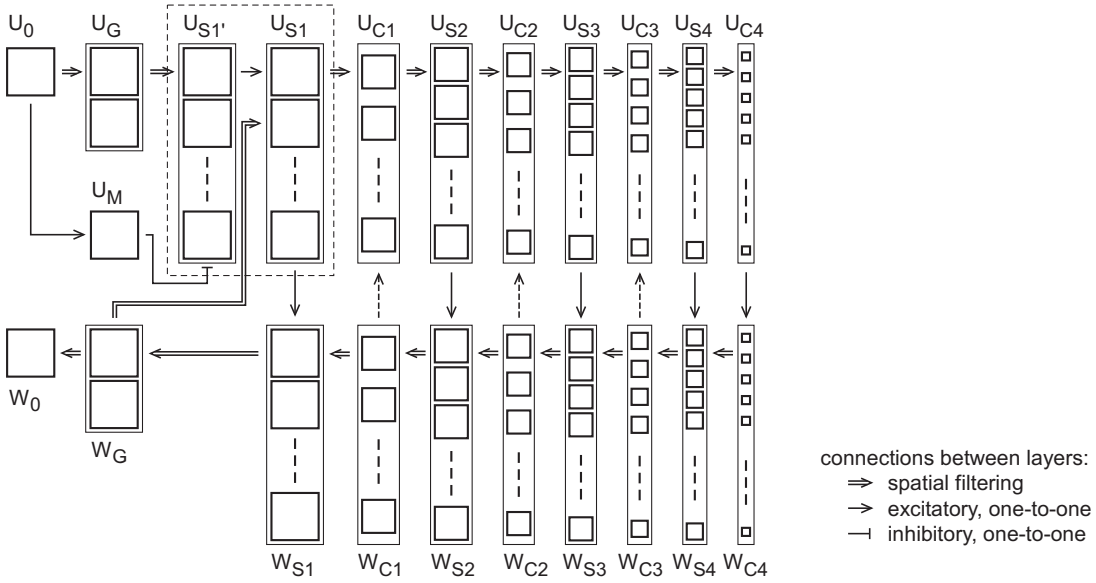


Figure 21. The architecture of a neural network model that can recognize and restore partly occluded patterns.

4.2 Forward Paths

The forward paths mainly take charge of recognizing partly occluded patterns [26]. As has been discussed above, local features extracted near occluded objects have a large tendency of being irrelevant to an occluded pattern. Hence the transmission of irrelevant features toward higher stages is blocked, by suppressing the response of feature-extracting cells near the occluding objects, as illustrated in Figure 22.

We will discuss the network in more detail. The architecture and the function of the forward paths (from U_0 through U_{C4} in Figure 21) is almost the same as that of the neocognitron [12][13][14].

Layer U_0 is the input layer consisting of photoreceptor cells, to which visual pattern is presented. Layer of contrast-extracting cells (U_G) follows U_0. The layer consists of two cell-planes: one with concentric on-center receptive fields, and the other with off-center receptive fields. The output of U_G is sent to the S-cell layer of the first stage (U_{S1}). The present model has four stages of S-cell and C-cell layers (U_{Sl} and U_{Cl}). S-cells have modifiable input connections, which are determined by the learning. They work as feature-extracting cells after finishing the learning. Input connections to C-cells are fixed and not modified by the learning. The response of each cell-plane of an S-cell layer U_{Sl} is spatially blurred in the corresponding cell-plane of the succeeding C-cell layer U_{Cl}. Layer U_{C4}, which is in the highest stage of the forward paths, is the recognition layer, whose response shows the final result of pattern recognition by the network.

A layer U_M, which is called a *masker layer*, is added to this neocognitron-like network [26]. The masker layer, which has only one cell-plane, detects and responds only to occluding objects. The shape of the occluding objects is detected and appears in U_M, in the same shape and at the same location as in the input layer U_0.[4] There are topographically ordered and slightly diverging inhibitory connections from layer U_M to all cell-planes of U_{S1} (to be more exact, $U_{S1'}$, which will be explained in Section 4.3 below). U_{S1} is the S-cell layer of the first stage,

[4]The process of segmenting occluding objects is not perfect yet in the present model, and occluding objects are detected based on the difference in brightness between occluding objects and occluded patterns.

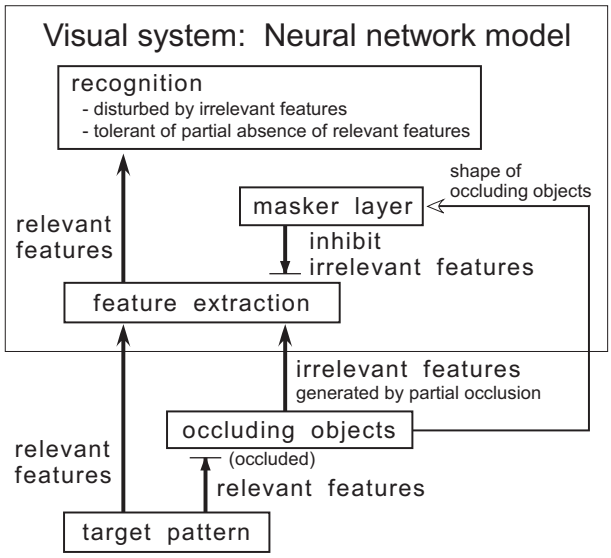


Figure 22: The main stream of information in the forward paths of the model when recognizing a partly occluded pattern.

which consists of edge-extracting cells. The inhibitory signals from layer U_M suppress the responses to features irrelevant to the occluded pattern. Hence only local features relevant to the occluded pattern are transmitted to higher stages of the network. Since the neocognitron, like human beings, has some tolerance for the absence of local features, the target pattern can be recognized correctly if irrelevant features do not reach higher stages.[5]

Figure 23 shows how the edges are extracted by U_{S1} and suppressed by the inhibition from U_M. Figure 23(a) is a stimulus presented to U_0. The sum of the responses of all cell-planes of U_{S1}, before and after being inhibited by U_M, are shown in (b) and (c), respectively. Namely, they show edges of all orientations that are extracted by U_{S1}. We can roughly express that our model analyzes the information of (c), and not (b), to recognize and restore the occluded pattern (a). If we express this using terminology of border ownership [27], this is an operation to ignore contours owned by occluding objects.

(a) (b) (c)

Figure 23: Edges extracted by U_{S1} from (a) the partly occluded pattern, (b) before and (c) after receiving inhibition from U_M.

4.3 Backward Paths

The cells in the backward (top-down) paths are arranged in the network making a mirror image of the cells in the forward (bottom-up) paths. Most of the connections also make a mirror image of those of the forward paths. The input connections diverging backward from S-cells of W_{Sl}, which are input connections to C-cells of W_{Cl-1}, have the same strength as the connections converging to S-cells of U_{Sl} in the forward paths. Only the direction of signal flow through the connections is reversed in the backward paths.

This architecture resembles that of the *Selective Attention Model*, which the author proposed previously [25]. This previous model already has a function of restoring damaged patterns. In the previous model, however, the

[5] A network with this mechanism can recognize all occluded patterns in Figure 19(b) correctly as 'A', 'B', $\cdots$ and 'F', where occluding objects are visible. Results of recognizing patterns in Figure 19(a), where occluding objects are not visible, depend on the set of learned patterns. In the case of our previous model [26], which has only forward paths and lacks the backward paths, patterns in Figure 19(a) were recognized as 'I', 'E', (rejected), 'E', 'L' and 'G', respectively.

backward signals are fed back only from the recognition cells at the highest stage of the forward paths. If an unlearned novel pattern is presented to the input layer of the network, no recognition cell will respond at the highest stage, hence top-down signals cannot start flowing.

In the present model, the backward signals are not only fed from U_{C4} to W_{C4} at the highest stage, but also from U_{Sl} to W_{Sl} at every intermediate stages ($l = 1, 2, 3$).

By the introduction of blurring operations, the forward paths, like the neocognitron, acquire an ability to recognize patterns robustly without being affected so much by deformation or shift in location. The blurring operations are applied, not only spatially, but also among resembling features. Blurring among resembling features are automatically performed by simply decreasing the thresholds of feature-extracting S-cells [28].

Although the blur in space and among resembling features in the forward paths are very useful for robust recognition of deformed patterns, some deblurring operations are required in the backward paths to restore a sharp image of partly occluded patterns.

In our model, backward signals from W_{Cl} and the signals from the forward layer U_{Sl} are simply added at W_{Sl} of the backward paths, where the former signals are usually much blurred than the latter signals. If both signals are transmitted to the lowest stage of the network without mutual interaction, a blurred image would appear strongly in W_0, and a clear restoration of the pattern could not be expected. It is desired that, in the backward paths of the network, the sharper signals from U_{Sl} predominate over the blurred signals from W_{Cl} in the locations where both signals exist, and the blurred signals are used only in the places where sharper signals are missing. This desired operation is automatically performed in two steps: by the backward connections from W_{Sl} to W_{Cl-1}, and by the backward connections from W_{Cl-1} to W_{Sl-1}. The former connections are used mainly for spatial deblurring, and the latter mainly for deblurring among resembling features. Incidentally, the latter connections are modified during the learning phase, using the responses of the forward cells [24].

Layer W_G in the backward paths corresponds to U_G in the forward paths. The positive and negative contrast components of a restored pattern are expected to appear here.

Layer W_0 is a virtual layer, which is utilized to monitor how the occluded pattern is *perceived* by the model. The model try to restore the original pattern in W_0 from the response of layer W_G. The restoration is done by diffusing the response of W_G. On-center cells of W_G send positive signals to W_0, and off-center cells send negative signals. These signals are diffused to the neighboring cells, but the signal of the opposite polarity works as a barrier to the diffusion.

Layer U_{S1} in the forward paths receives signals from W_G, as well as from U_G via $U_{S1'}$, and extract oriented edges. The sum of these two signals appears in U_{S1}. The connections from W_G to U_{S1} are the same as those from U_G to $U_{S1'}$. The reason why a separate layer $U_{S1'}$ is placed only in the forward paths is that the inhibitory signals from masker layer U_M should suppress only signals from U_G, and should not affect the signals from W_G of the backward paths.

Multiple positive feedback loops are thus generated in the network by the forward and the backward paths. The response of the network gradually changes while signals circulate in the feedback loops.

4.4 Restoration of Deformed Versions of Learned Patterns

The model was simulated on a computer. The network was initially trained to recognize alphabetical characters of a font. To demonstrate that the model can accept deformed patterns that have not been shown during the training phase, alphabetical characters of a different font are used for the test after having finished learning.

We test how partly occluded patterns are recognized and restored. Test patterns before occlusion are all deformed versions of the training patterns.

Figure 24 shows an example of the response of the network that has finished learning. Partly occluded pattern 'A' is presented to the input layer U_0. The response of the recognition layer U_{C4} shows that the occluded pattern is correctly recognized as 'A'. The shape of the restored pattern, which is *perceiver* by the model, appears in W_0.

Figure 25 summarizes how the network has restored, or *perceived*, partly occluded patterns. In each pair of patterns, stimulus given to input layer U_0 and the *perceived* pattern appeared in W_0 are shown.

Similarly to human perception, the model responds differently to patterns (b) of Figure 20, in which the black patterns are identical but different gray objects are placed. The left and right patterns are recognized as 'R' and 'B', respectively, and the patterns that the model *perceives* are shown in the second row of Figure 25. Similarly, left and right patterns of Figure 20(a) are recognized as 'C' and 'O', respectively, and the patterns that the model *perceives* are shown in the bottom row of Figure 25. The third row of Figure 25 shows another example, where the

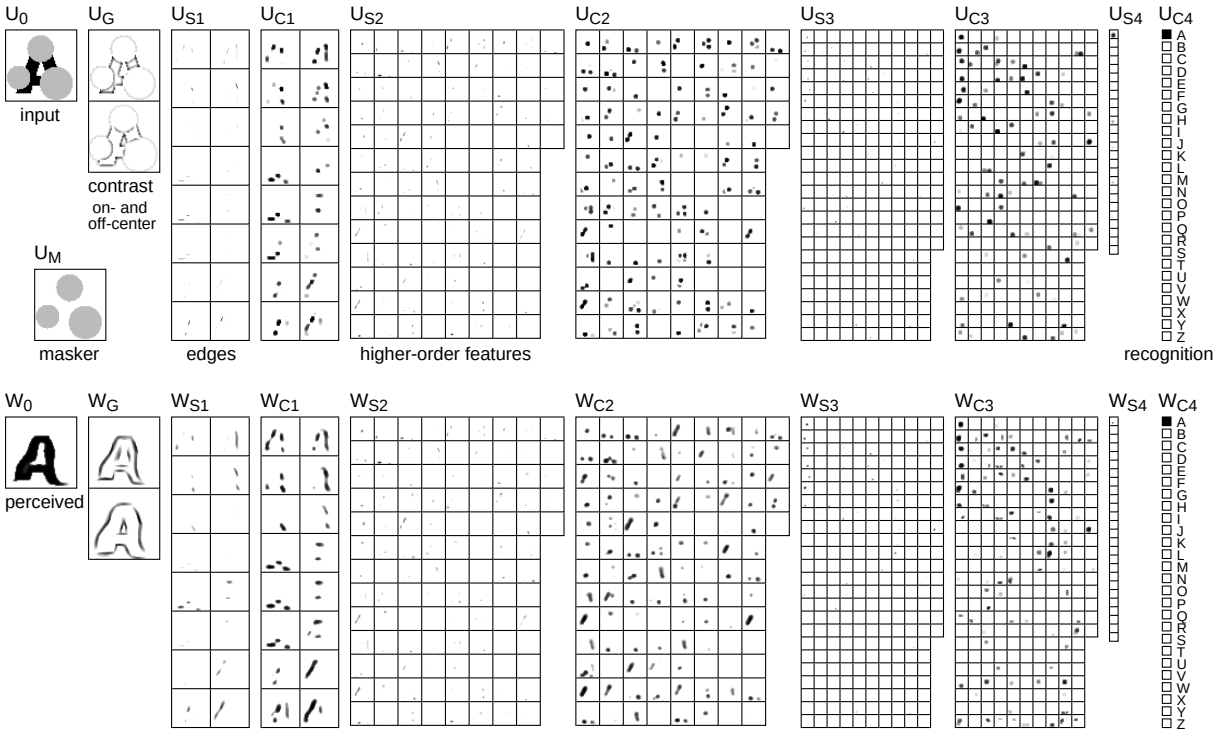


Figure 24: An example of the response of the network that has finished learning. Input pattern U_0 is recognized as 'A' at U_{C4}, and is restored in W_0.

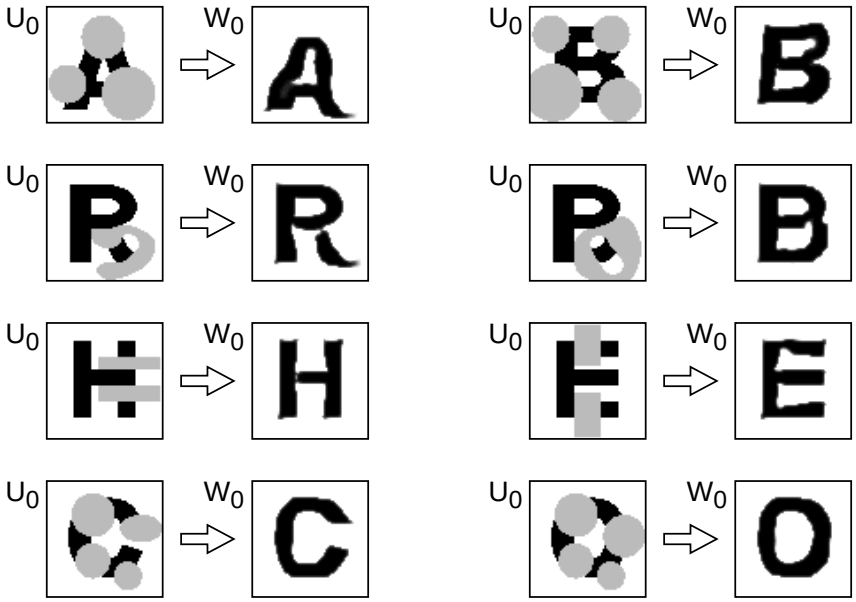


Figure 25. The responses of W_0, which are restored from partly occluded patterns presented to U_0.

same black object is perceived as 'H' and 'E', depending on the different placement of gray objects.

As can be seen in Figure 25, the restored patterns are not necessarily identical in shape to the original patterns before occlusion. It is natural, however, that the model, which recognizes patterns accepting some deformation, cannot imagine the exact shapes in the occluded areas because it has not seen the exact shapes of the patterns. It is noteworthy that, despite some deformation, all essential components of patterns have been correctly reconstructed.

4.5 Restoration of Novel Patterns

To emulate a response to a novel pattern that has not been learned yet, the output of layer U_{S4} is forced to be suppressed. This is an emulation of a situation, in which the model has already learned varieties of local features from various learning patterns but has not learned the global shape of the target pattern, which is partly occluded.

Figure 26 compares how learned and novel patterns are restored. It can be seen that the completion of missing portion of an novel target pattern is made mainly by interpolating and extrapolating contours of the unoccluded parts of the pattern, rather than by knowledge on the global shape of the target pattern. We can also see from the figure that resemblance of local features to other learned patterns are also utilized for the restoration.

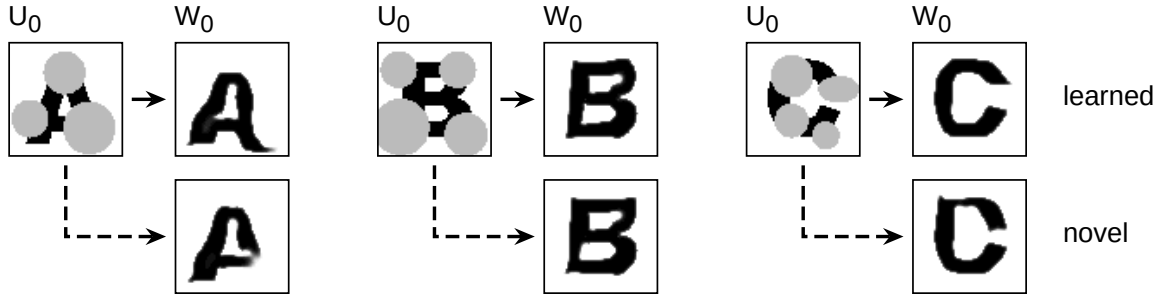


Figure 26. Comparison of restoration from learned and novel patterns.

5. Conclusion

This review has introduced several results of modeling neural networks for visual information processing. To extend modeling researches further, we first have to deepen our knowledge on the mechanism of the biological brain. The neural network of the brain is a system that works under several constraints. The cost-minimizing principle, by which the second boom of the neural network research was triggered, is one of such constraints, but there must be many other important constraints. We now have come to the stage to uncover new principles that control the neural networks of the brain. It is a very attractive research area, that lead to, not only explication of the mechanism of the brain, but also invention of design principles of intelligent information processing systems of the next generation.

Acknowledgement

This work was partially supported by Grants-in-Aid for Scientific Research #14380169 from Japan Society of Promotion of Science.

References

[1] H. Saito, M. Yukie, K. Tanaka, K. Hikosaka, Y. Fukada, E. Iwai: “Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey”, *J. Neuroscience*, **6**, 145–157, 1986.

[2] K. Tanaka, H. Saito: “Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey”, *J. Neurophysiology*, **62**, 626–641, 1989.

[3] K. Tanaka, Y. Fukada, H. Saito: “Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey”, *J. Neurophysiology*, **62**, 642–656, 1989.

[4] C. J. Duffy, R. H. Wurtz: “Sensitivity of MST neurons to optic flow stimuli. I. a continuum of response selectivity to large-field stimuli”, *J. Neurophysiology*, **65**, 1329–1345, 1991.

[5] M. S. A. Graziano, R. A. Andersen, R. J. Snowden: “Tuning of MST neurons to spiral motions”, *J. Neuroscience*, **14**, 54–67, 1994.

[6] C. J. Duffy, R. H. Wurtz: "Sensitivity of MST neurons to optic flow stimuli. II. mechanisms of response selectivity revealed by small-field stimuli", *J. Neurophysiology*, **65**, 1346–1359, 1991.

[7] J. J. Koenderink: "Optic flow", *Vision Research*, **26**, 161–179, 1986.

[8] K. Tohyama, K. Fukushima: "Neural network model for extracting optic flow", *Neural Networks*, **18**[5/6], pp. 549–556, 2005.

[9] K. Tanaka, K. Hikosaka, H. Saito, M. Yukie, Y. Fukada, E. Iwai: "Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey", *J. Neuroscience*, **6**, 134–144, 1986.

[10] D.-K. Xiao, S. Raiguel, V. Marcar, G. A. Orban: "The spatial distribution of antagonistic surround of MT/V5 neurons", *Cerebral Cortex*, **7**, 662–677, 1997.

[11] K. Fukushima: "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biological Cybernetics*, **36**[4], pp. 193–202, 1980.

[12] K. Fukushima: "Neocognitron: A hierarchical neural network capable of visual pattern recognition", *Neural Networks*, **1**[2], pp. 119–130, 1988.

[13] K. Fukushima: "Neocognitron for handwritten digit recognition", *Neurocomputing*, **51**, pp. 161–180, 2003.

[14] Visiome Platform, from which computer programs for several types of neocognitron can be downloaded: http://platform.visiome.org/index.html

[15] C. A. Burbeck, S. M. Pizer: "Object representation by cores: identifying and representing primitive spatial regions", *Vision Research*, Vol. 35, No. 13, pp. 1917–1930, 1995.

[16] S. Sato, S. Miyake: "A model of visual attention with Gestalt principles on shapes of objects" (in Japanese), *Technical report of IEICE*, No. NC2002-186, pp. 79–84, 2003.

[17] S. C. Dakin, J. Watt: "Detection of bilateral symmetry using spatial filters", in: C. W. Tyler, ed., *Human Symmetry Perception and its Computational Analysis*, Lawrence Erlbaum Associates, New Jersey, 2002, pp. 187–207.

[18] D. E. Rummelhart, G. E. Hinton, R. J. Williams: "Learning internal representations by error propagation", in: D. E. Rummelhart, J. L. McClelland, PDP Research Group, eds., *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, Bradford Book, MIT Press, Cambridge, MA, 1986, pp. 318–362.

[19] T. J. Sejnowski, P. K. Kienker, G. E. Hinton: "Learning symmetry groups with hidden units: beyond the perception", *Physica*, Vol. 22D, pp. 260–275, 1986.

[20] C. Latimer, W. Young, C. Stevens: "Modelling symmetry detection with back-propagation network", in: C. W. Tyler, ed., *Human Symmetry Perception and its Computational Analysis*, Lawrence Erlbaum Associates, New Jersey, 2002, pp. 209–225.

[21] I. Kovács, Á. Fehér, B. Julesz: "Medial-point description of shape: a representation for action coding and its psychological correlates", *Vision Research*, Vol. 38, pp. 2323–2333, 1998.

[22] G. V. Tonder, Y. Ejima: "The patchwork engine: image segmentation from shape symmetries", *Neural Networks*, Vol 13, pp. 291–303, 2000.

[23] K. Fukushima: "Use of non-uniform spatial blur for image comparison: Symmetry axis extraction", *Neural Networks*, **18**[1], pp. 23–32, 2005.

[24] K. Fukushima: "Restoring partly occluded patterns: a neural network model", *Neural Networks*, **18**[1], pp. 33–43, 2005.

[25] K. Fukushima: "Neural network model for selective attention in visual pattern recognition and associative recall", *Applied Optics*, **26**[23], pp. 4985–4992, 1987.

[26] K. Fukushima: "Recognition of partly occluded patterns: a neural network model", *Biological Cybernetics*, **84**[4], pp. 251–259, 2001.

[27] H. Zhou, H. S. Friedman, R. von der Heydt: "Coding of border ownership in monkey visual cortex", *Journal of Neuroscience*, Vol. 20, No. 17, pp. 6594–6611, 2000.

[28] K. Fukushima, M. Tanigawa: "Use of different thresholds in learning and recognition", *Neurocomputing*, **11**[1], pp. 1–17, 1996.

Kunihiko Fukushima is a Visiting Professor at the Graduate School of Informatics, Kansai University, Takatsuki, Osaka, Japan. He received a B.Eng. degree in electronics in 1958 and a PhD degree in electrical engineering in 1966 from Kyoto University, Japan. He was a professor at Osaka University from 1989 to 1999, at the University of Electro-Communications from 1999 to 2001, and at Tokyo University of Technology from 2001 to March 2006. Prior to his Professorship, he was a Senior Research Scientist at the NHK Science and Technical Research Laboratories. His special interests lie in modeling neural networks of the higher brain functions, especially the mechanism of the visual system. He invented the “neocognitron” for deformation invariant pattern recognition. He was the founding President of the JNNS and a founding member on the Board of Governors of the INNS. (Homepage: http://www.teu.ac.jp/media/˜fukushima/)