Symmetry Axis Extraction by a Neural Network

Kunihiko Fukushima, Masayuki Kikuchi

Tokyo University of Technology
1404-1, Katakura, Hachioji, Tokyo 192-0982, Japan
fukushima@media.teu.ac.jp, kikuchi@cs.teu.ac.jp

Abstract

This paper proposes a neural network model that extracts axes of symmetry from visual patterns. The input patterns can be line drawings, plane figures or gray-scaled natural images taken by CCD cameras.

The model is a hierarchical multi-layered network, which consists of a contrast-extracting layer, edge-extracting layers (simple and complex types), and layers extracting symmetry axes. Its architecture resembles that of the lower stages of the neocognitron. The model extracts oriented edges from the input image first, and then tries to extract axes of symmetry.

To reduce the computational cost, the model checks conditions of symmetry, not directly from the oriented edges, but from a blurred version of them. The use of blurred signals endows the network with a large tolerance to deformation of input patterns, too. It is important to get blurred signals, not directly from the input image, but from the oriented edges. If the input image is directly blurred, most of the important features in the image will be lost. Since the model uses oriented edges, however, most of the important features can still remain even after the blurring operations, by which information of edge locations becomes ambiguous.

1 Introduction

Mirror-images, or bilateral symmetries, are perceptually salient. Symmetry about an axis usually attracts our eyes more strongly than other comparative regularities in a figure. This paper proposes a neural network model that extracts axes of symmetry.

Various models or algorithms have already 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 [1][2], or spatial filters [3] to make this comparison. Some others try to create such neural networks by learning [4][5][6]. 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. It is desired to reduce this computational cost.

Some of the methods try to extract medial axes or skeletons, which are mixtures of axes of global and local symmetries kovacs98[8]. 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,

The model that we propose extracts oriented edges from the input image first, and then tries to extract axes of symmetry. To reduce the computational cost, and to increase the robustness against noise and deformation of input patterns, the model checks conditions of symmetry, not directly from the oriented edges, but from a blurred version (low-resolution responses covering a large area) of them. Since this method checks a necessary condition, and not always a sufficient condition, of symmetry, there is a possibility of generating spurious outputs. To eliminate spurious outputs, the check of symmetry is performed, not at a single resolution, but at a small number of (three, for example) different resolutions.

2 Neural Network Model

2.1 Outline of the Network Architecture

The model is a multilayered network with a hierarchical architecture, which is illustrated in Fig. 1. It resembles the architecture of the lower stages of the neocognitron [9][10]. Each layer consists of a number of cell-planes. Incidentally, a cell-plane is a group of cells that are arranged retinotopically and share the same set of input connections. As a result, they all have receptive fields of an identical characteristic, but the locations of the receptive fields differ from cell to cell.

The model consists of a number of layers connected in an 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).

2.2 Contrast Extraction

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.

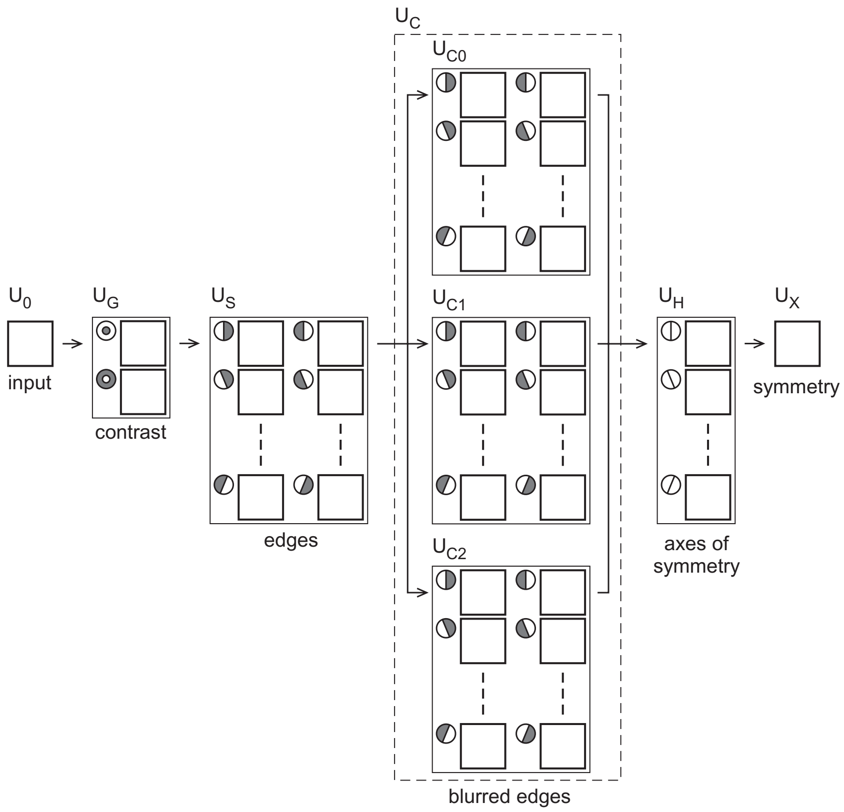

Figure 1: Architecture of the proposed network.

Cells of contrast-extracting layer U_G resemble retinal ganglion cells or lateral geniculate nucleus cells. Layer U_G consists of two cell-planes: one with concentric on-center receptive fields, and the other with off-center receptive fields. The former cells extract positive contrast in brightness, whereas the latter extract negative contrast from the image presented to the input layer.

To be more specific, the spatial distribution of the input connections to an on-center cell take the shape of a Mexican hat, and off-center cells have connections of opposite polarities. These connections are so designed that the total sum of the connections converging to a single cell is equal to zero. This means that the dc component of spatial frequency of the input pattern is eliminated in the contrast-extracting layer U_G. As a result, the output from layer U_G is zero in the area where the brightness of the input pattern is flat.

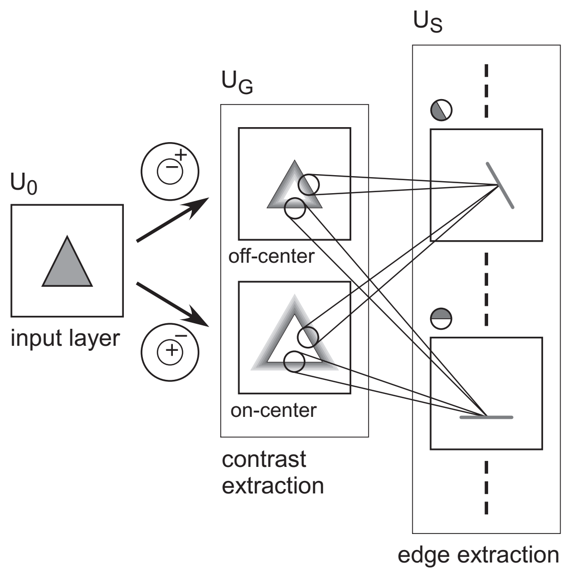

Figure 2: Extraction of oriented edges from positive and negative contrast components.

2.3 Edge-Extraction by S-cells

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. We will also call this layer simply S-cell layer. Each S-cell receives connections from both on- and off-center cells of U_G and extracts edges of a particular orientation, as illustrated in Fig. 2.

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 $2\pi/K = 11.25°$, and the preferred orientation of the kth cell plane is $\alpha_k = 2\pi k/K$. Layer U_S thus decompose the input image into edge components of various orientations.

The input connections to S-cells are determined with the same process as the supervised learning for the neocognitron [10]. Once they have been initially trained, they are fixed afterward. To train a cell-plane, a "teacher" presents a training pattern, which is a straight edge of orientation $\alpha_k = 2\pi k/K$, to the input layer of the network. The teacher then chooses from the kth cell-plane a cell whose receptive field center is located at a point on the edge. (Any cell located on the edge is enough for this purpose.) The cell takes the place of the seed cell of the cell-plane and has its input connections strengthened. The amount of increment of each connection is proportional to the output of the presynaptic cell. The speed of increasing the connections is set so large that the training of a seed cell is completed by only a single presentation of each training pattern. Because the connections are shared by all cells in the cell-plane, the training of input connections of all cells in the cell-plane finishes with this single process. Cells of the cell-plane thus come to respond selectively to edges of the orientation of the training pattern.

The selectivity of an S-cell in extracting edges can be controlled by the value of its threshold. In our model, the threshold is set low enough for S-cells to accept edges of slightly different orientations.

2.4 Blurring by C-cell Layers

The output of layer U_S is fed to layer U_C (edge-extracting layer of complex-type), where a blurred version of the response of layer U_S is generated. Layer U_C consists of C-cells, which resemble complex cells in the primary visual cortex. We will also call this layer simply C-cell layer. Connections from S-cells to C-cells resemble the classical hypothesis by Hubel and Wiesel.

Layer U_C consists of M sub-layers, U_{Cm}, each of which have a different degree of blur. These M sub-layers are arranged in parallel in the network, and each sub-layer has the same number (K) of cell-planes as U_S. In the computer simulation, we will take $M=3$.

Each C-cell of the mth sub-layer has input connections of a bell-shaped spatial distribution of radius A_{Cm}. It integrates the responses of a group of S-cells of the corresponding cell-plane whose receptive field-centers are located within an area of radius A_{Cm}. A blurred version of the response of layer U_S thus appears in layer U_C.

In the computer simulation, radiuses A_{Cm} for different sub-layers are adjusted in such a way that $A_{Cm} = 2^m A_{C0}$ holds approximately.

2.5 Extraction of Axes of Symmetry

2.5.1 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 model) on one side of the axis always has its counterpart on the other side of the axis as shown in Fig. 3.

We will check if a point, o, is on the axis of symmetry of orientation α_k. 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 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.

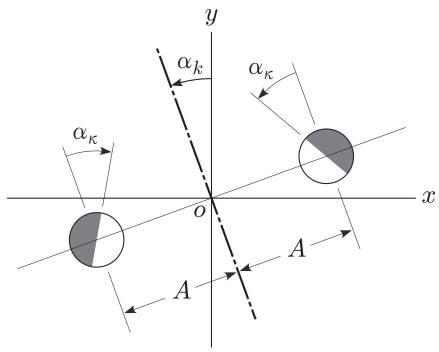

Figure 3: Principles of extracting an axis of symmetry.

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 model, not directly from the output of edge-extracting layer U_S, but from the output of sub-layers of U_C. Incidentally, each C-cell (cell of U_C) sums the response of U_S using bell-shaped input connections of radius A_C. We calculate h only at a pair of locations that satisfy $A = A_C$, and sum up h for all orientations of edges, α_κ. 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, and not always a sufficient, condition of symmetry. In other words, there is a possibility of spurious outputs at places that is not on an axis of symmetry.

In our model, we calculate H with $M(=3)$ different values (resolutions) of A_C and sum them up. 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 usually be suppressed by mutual interaction between M different resolutions.

It is important to note here that the shape of input connections to a C-cell, through which the responses of S-cells are gathered, should be of a bell-shape, and should not be flat like a cylinder. It is required that the response of a C-cell largely changes by a shift in location of an edge. In other words, a signal of a certain strength from a differ-

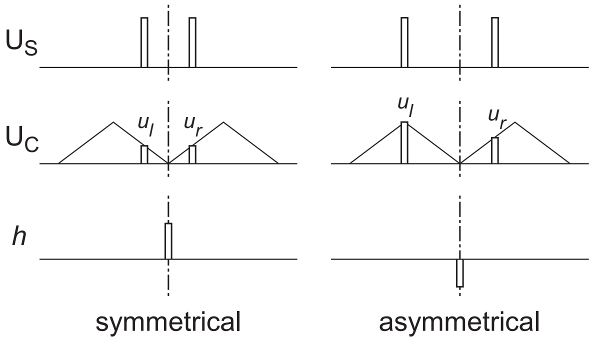

Figure 4: Illustration of the process of calculating h. When the stimulus is symmetrical, $u_r = u_l$ holds. Hence, h takes a positive value. When the stimulus is asymmetrical, however, u_r and u_l take different values, and h becomes negative even though the values of u_r and u_l themselves are large.

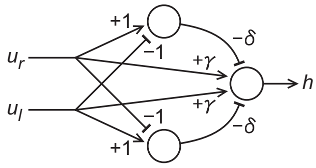

Figure 5: An example of a circuit calculating $h = \gamma(u_r + u_l) - \delta|u_r - u_l|$.

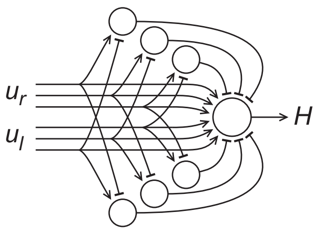

Figure 6: An example of a circuit for an H-cell, which calculates H.

ent location of U_S should elicit a different response from the C-cell. 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. This is illustrated in Fig. 4.

Incidentally, calculation of h can easily be done by a simple neural circuit, like the one shown in Fig. 5, for example. The circuit 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.

Circuit for calculating H can be easily made by combining this circuit. Since the summation of h is made without threshold operations, the cell on the right side of Fig. 5 has to be used in common for all calculations of h, and the circuit for H becomes like the one shown in Fig. 6.

2.5.2 Layers Extracting Symmetry Axis

In the proposed model, symmetry-axis-extracting layer U_H follows layer U_C. Each cell (H-cell) of layer U_H analyzes the response of U_C and calculates H, which is a weighted sum of h for different edge-orientations and for different resolutions, and checks if its receptive-field center is on an axis of symmetry.

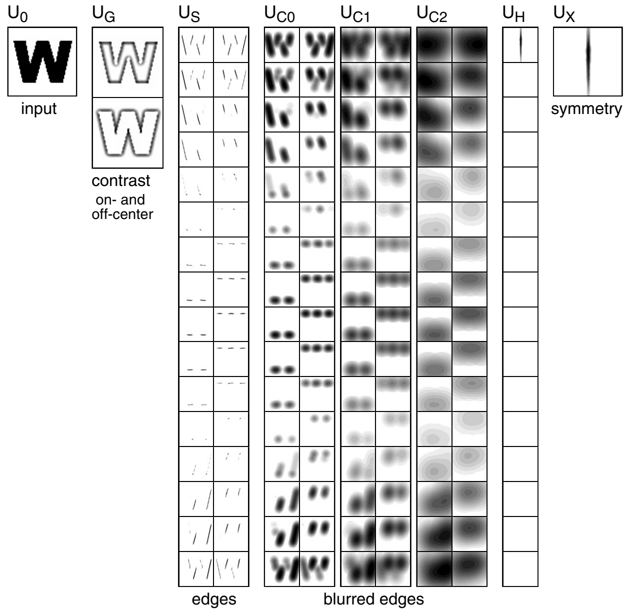

Figure 7: An example of the response of the network.

Layer U_H consists of $K/2$ cell-planes depending on the orientations of axes of symmetry. Each cell in a cell-plane checks if its receptive-field center is on an axis of symmetry of a particular orientation. It should be noted here that the number of cell-planes is not K but $K/2$, because the orientations of axes of symmetry are distributed within the range of 180°, while the orientations of edges within the range of 360°. The principles of extracting symmetry by U_H will be discussed below in more detail.

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.

3 Computer Simulation

The model is simulated on a computer. Fig. 7 shows how the cells in the proposed network respond to a W-shaped pattern.

Layers U_0 and U_{Cm} are slightly larger in size than those displayed in the figure. In cell-planes of U_{Cm}, the blurred response of U_S might spatially spread across the boundaries of the cell-planes. If these leaked responses are treated as zero, spurious responses might appear in the peripheral part of the cell-planes of U_H. Hence the cell-planes of U_{Cm} are designed to be large enough to prevent the generation of such spurious responses, but Fig. 7 shows only their responses from the central areas of the same size as those of U_S.

Similarly, input layer U_0 actually is slightly larger than that displayed in Fig. 7. This is required when half-toned

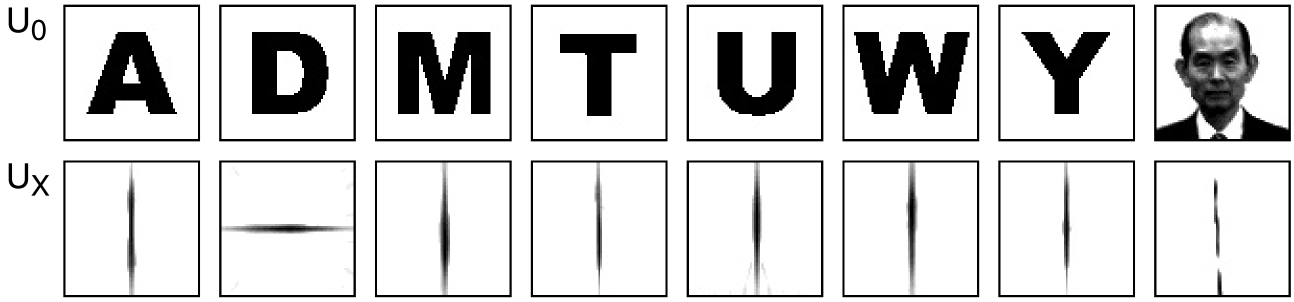

Figure 8: Axes of symmetry extracted from various input patterns.

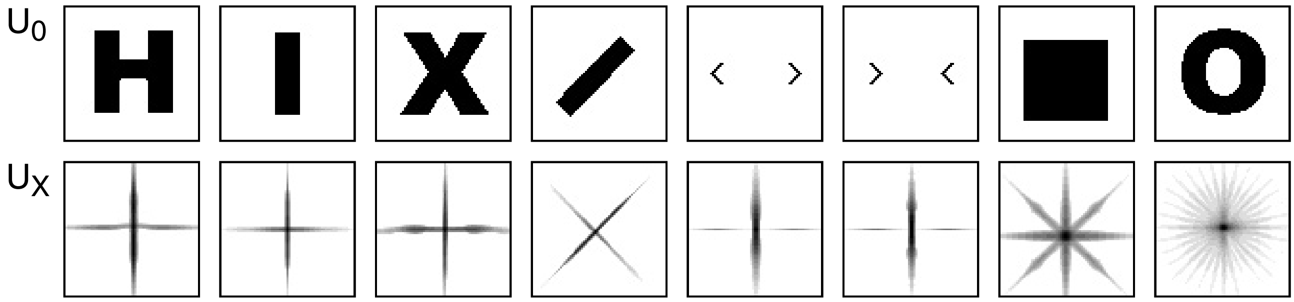

Figure 9: Axes of symmetry extracted from patterns having two or more axes.

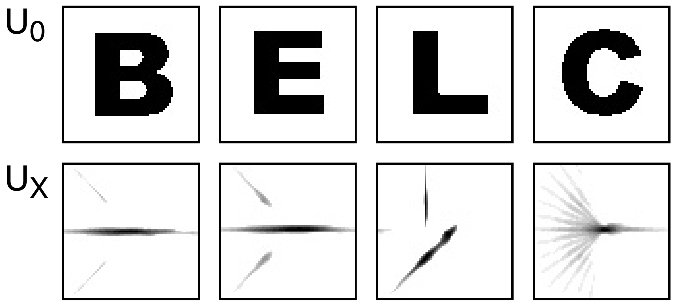

Figure 10: Some example of patterns, from which axes of local symmetry are also extracted together with global symmetries.

images are presented to U_0. If U_0 and U_G are of the same size, and if the images outside of the layer U_0 are treated as zero, contrasts of brightness are erroneously extracted at the border of the layer. To suppress the generation of such erroneous responses, layer U_0 is actually larger than U_G by the radius of the connections A_G in four sides of the layer. Fig. 7 displays, however, only the responses from the central areas of the same size as those of U_S.

Figs. 8–10 summarizes the responses of symmetry-extracting layer U_X for various input patterns given to U_0. As can be seen from Fig. 8, symmetry axes are extracted correctly even from complicated figures. The model can extract symmetry axis even from a photograph of a human face, which is taken by a CCD camera and contains small amount of asymmetry.

Even when an input pattern has two or more symmetry axes, they all can be extracted as shown in Fig. 9. It can be seen from Figs. 8 and 9, small amount of asymmetry caused by deformation of the pattern can be accepted. This tolerance is mainly attained with the blurring operation by U_C.

From pattern ‘O’, which is a vertically elongated circle, symmetry axes of all orientations are extracted, because the vertical elongation is not so large as to completely destroy the symmetry about slanted axes. Among these extracted axes, however, vertical and horizontal axes, which are the axes of strict symmetry, are stronger than the other axes.

In some patterns shown in Fig. 10, axes of local symmetry are extracted (although they are not so strong) together with an axis of global symmetry. To suppress extracting such local symmetry-axes, another sub-layer with a larger value of A_C is required in U_C. In other words, U_C will require a more number of sub-layers ($M = 4$, for example). A number of axes of local symmetry extracted in the left side of pattern ‘C’ will also be suppressed by introducing another sub-layer.

4 Discussions

This paper proposed a neural network model that extracts axes of symmetry. The model extracts oriented edges (at layer U_S) from an input image first, and then tries to extract axes of symmetry. The model checks conditions of symmetry, not directly from the oriented edges (response of layer U_S), but from the response of layer U_C, which is a blurred version of the response of layer U_S. Layer U_C consists of C-cells, which have larger receptive fields and yield low-resolution responses covering a large area of the input layer. The use of C-cells has increased the tolerance for deformation of the input pattern, and greatly reduced computational costs for extracting symmetry axes.

Use of oriented edges (U_S), and not the input image itself (U_0 or U_G), is also useful, when comparison is made for 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.

Acknowledgment

This work was partially supported by Grants-in-Aid-for-Scientific-Research #14380169 and #14780257, and Special Coordination Fund for Promoting Science and Technology (Project on Neuroinformatics Research in Vision), both from the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

References

[1] 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.

[2] 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.

[3] 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, pp. 187–207, 2002.

[4] 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, pp. 318–362, 1986.

[5] 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.

[6] 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, pp. 209–225, 2002.

[7] 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.

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

[9] K. Fukushima: "Neocognitron: A hierarchical neural network capable of visual pattern recognition", *Neural Networks*, vol. 1, no. 2, pp. 119–130, 1988.

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