Neural Networks, Vol. 2, pp. 413–420, 1989
Printed in the USA. All rights reserved.

0893-6080/89 $3.00 + .00

ORIGINAL CONTRIBUTION

Analysis of the Process of Visual Pattern Recognition by the Neocognitron

KUNIHIKO FUKUSHIMA

Osaka University

(*Received* 2 *December* 1988; *revised and accepted* 5 *March* 1989)

Abstract—*A neural network model of visual pattern recognition called the "neocognitron," was earlier proposed by the author. It is capable of deformation-invariant visual pattern recognition. After learning, it can recognize input patterns without being affected by deformation, changes in size, or shifts in position. This paper offers a mathematical analysis of the process of visual pattern recognition by the neocognitron. The neocognitron is a hierarchical multilayered network. Its initial stage is an input layer, and each succeeding stage has a layer of "S-cells" followed by a layer of "C-cells." Thus, in the whole network, layers of S-cells and C-cells are arranged alternately. The process of feature extraction by an S-cell is analyzed mathematically in this paper, and the role of the C-cells in deformation-invariant pattern recognition is discussed.*

Keywords—Visual pattern recognition, Learning, Multilayered network, Deformation invariant, Neocognitron, Analysis.

1. INTRODUCTION

A neural network model of visual pattern recognition, called the "neocognitron," was earlier proposed by the author (Fukushima, 1980, 1988b; Fukushima & Miyake, 1982). The neocognitron is a hierarchical multilayered neural network capable of deformation-invariant visual pattern recognition. After learning, it can recognize input patterns without being affected by deformation, changes in size, or shifts in position. Even if the input pattern is deformed in shape, only one cell, corresponding to the category of the input pattern, is activated in the highest stage of the network. Other cells respond to patterns of other categories. This situation is illustrated in Figure 1.

We offer a mathematical analysis of the process of visual pattern recognition of the neocognitron model in this paper.

2. STRUCTURE AND BEHAVIOR OF THE NEOCOGNITRON

In the visual area of the cerebrum, neurons respond selectively to local features of a visual pattern, such as lines or edges in particular orientations. In an area higher than the visual cortex, cells exist that respond selectively to certain figures like circles, triangles or squares, or even human faces. Thus, the visual system seems to be a hierarchical structure, in which simple features are first extracted from a stimulus pattern, then integrated into more complicated ones. In this hierarchy, a cell in a higher stage generally receives signals from a wider area of the retina, and is more insensitive to the position of the stimulus. This kind of physiological evidence suggested a network structure for the neocognitron.

The neocognitron has a multilayered structure like Figure 2, in which each rectangle represents a two-dimensional array of cells.

The initial stage of the hierarchical network is an input layer, consisting of a two-dimensional array of receptor cells. Each succeeding stage has a layer consisting of cells called "S-cells" followed by another layer consisting of cells called "C-cells." Thus, in the whole network, layers of S-cells and C-cells are arranged alternately. The layer of C-cells at the highest stage is the recognition layer: the response of the cells in this layer is the final result of pattern recognition by the neocognitron.

S-cells are feature-extracting cells. Connections converging to these cells are variable, and may be reinforced by learning (or training). After learning, S-cells can extract features from input patterns. In other words, an S-cell is activated only when a particular feature is presented at a certain position to the input layer. The features extracted by the S-cells are determined during the learning process. Gen-

Requests for reprints should be sent to Professor Kunihiko Fukushima, Osaka University, Biophysical Engineering Dept., Toyonaka, Osaka 560, Japan.

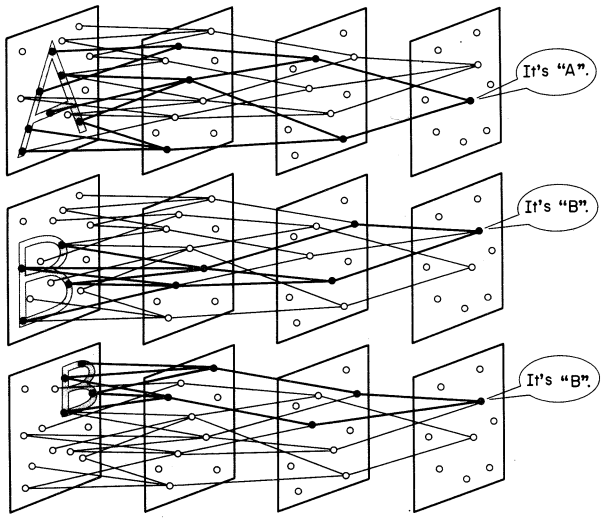

FIGURE 1. An explanation of how recognition cells at the highest stage of the neocognitron respond.

erally speaking, local features, such as lines in particular orientations, are extracted in the lower stages. More "global" features, such as parts of a training pattern, are extracted in higher stages. The process of learning and the mechanism of feature extraction by the S-cells are discussed in section 3.

C-cells are inserted in the network to allow for positional errors in the features of the stimulus. The connections from S-cells to C-cells are fixed and invariable.

As shown in Figure 2, each layer of S-cells or C-cells is divided into subgroups according to the feature to which they respond. The cells in each subgroup are arranged in a two-dimensional array. The connections converging to the cells in a subgroup are homogeneous: all the cells in a subgroup receive input connections of the same spatial distribution, in which only the position of the preceding cells shifts in parallel with the position of the cells in the subgroup. This condition of homogeneity holds for fixed connections and for variable connections. As discussed in section 5, the reinforcement of variable connections is always done under this condition. Such a subgroup of cells is called a "cell-plane."

Each C-cell receives signals from a group of S-cells that extract the same feature, but from slightly different positions. The C-cell is activated if at least one of these S-cells is active. Even if the stimulus feature is shifted in position and another S-cell is activated instead of the first one, the same C-cell keeps responding. Hence, the C-cell's response is less sensitive to shifts in the position of the input pattern.

In the whole network, with its alternate layers of S-cells and C-cells, the process of feature-extraction by the S-cells and toleration of positional shift by the C-cells is repeated. During this process, local features extracted in lower stages are gradually integrated into more "global" features. Figure 3 illustrates this situation schematically.

Finally, each C-cell of the recognition layer at the highest stage integrates all the information of the input pattern, and responds only to one specific pattern. In other words, only one C-cell, corresponding to the category of the input pattern, is activated. Other cells respond to the patterns of other categories.

Tolerating positional error a little at a time at each stage, rather than all in one step, plays an important role in endowing the network with the ability to recognize even distorted patterns. Figure 4 illustrates this situation. Let an S-cell in an intermediate stage

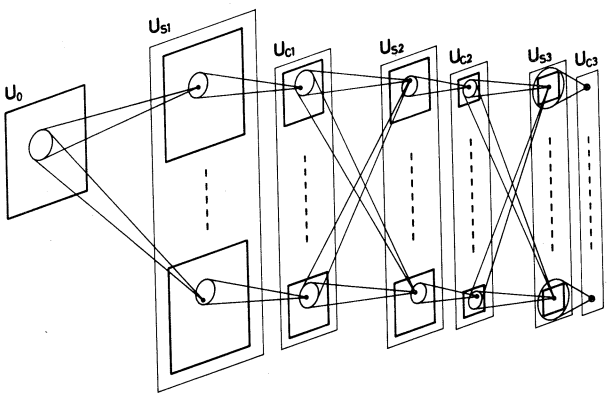

FIGURE 2. Hierarchical network structure of the neocognitron.

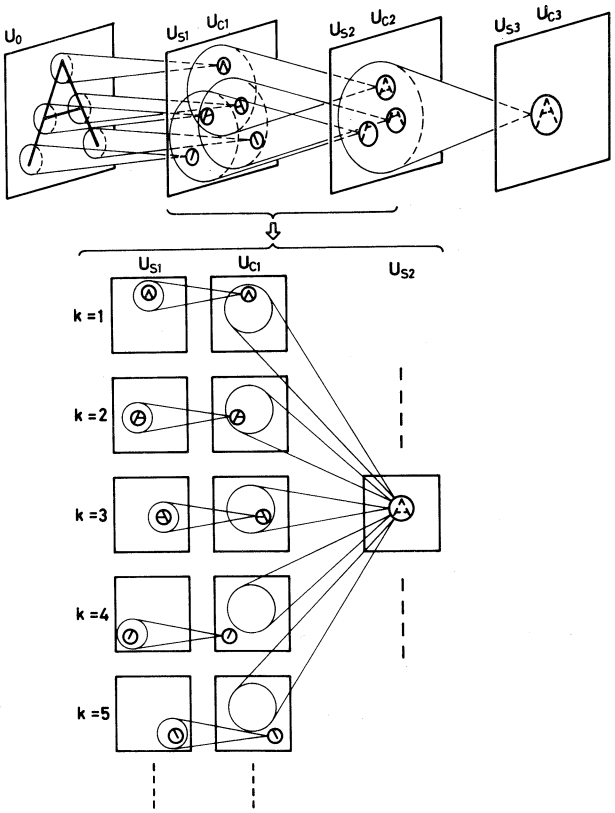

FIGURE 3. Illustration of the process of pattern recognition in the neocognitron (Fukushima, 1980).

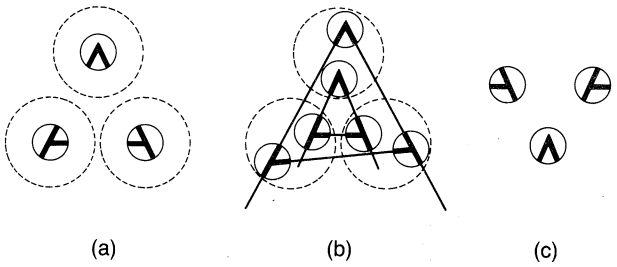

FIGURE 4. Illustration of the principle for recognizing deformed patterns (Fukushima, 1988a).

of the network have already been trained to extract a global feature consisting of three local features of a training pattern "A," as shown in Figure 4(a). The cell tolerates a positional error of each local feature if the deviation falls within the dotted circle. Hence, the S-cell responds to any of the deformed patterns shown in Figure 4(b). The toleration of positional errors should not be too large at this stage. If too large errors are tolerated at any one step, the network may come to respond erroneously, such as by recognizing a stimulus like Figure 4(c) as an "A" pattern.

Since errors in the relative position of local features are thus tolerated in the process of extracting and integrating features, the same C-cell responds in the recognition layer at the highest stage, even if the input pattern is deformed, changed in size, or shifted in position. The network recognizes the "shape" of the pattern independently of its size and position.

The principle of the neocognitron can be effectively used in various kinds of pattern recognition systems. For instance, a hand-written numeral recognition system has been developed with this principle (Fukushima, 1988b). Figure 5 shows some examples of deformed patterns which the neocognitron recognized correctly.

A mathematical description of the whole network of the neocognitron is given in the Appendix. We will analyze mathematically the process of feature-extraction by the S-cells and discuss the role of C-cells in detail in the text below.

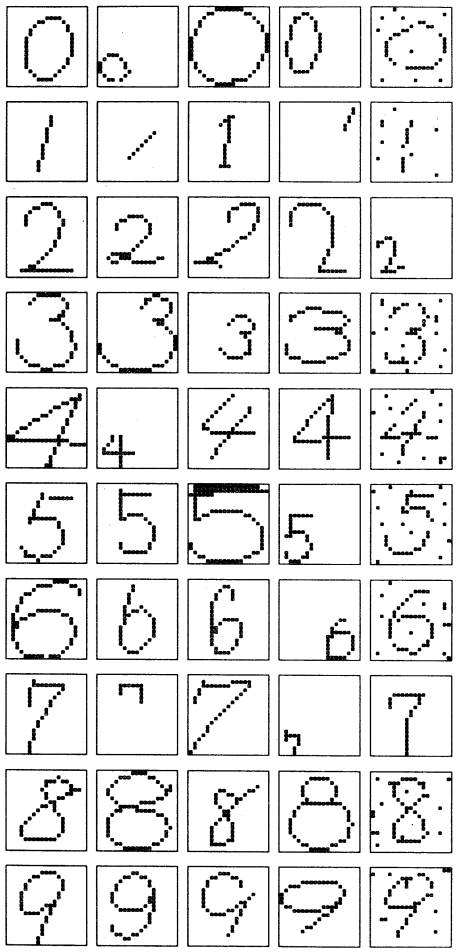

FIGURE 5. Some examples of deformed input patterns which the neocognitron has recognized correctly (Fukushima, 1988b).

3. FEATURE EXTRACTION BY AN S-CELL

3.1. Connections Converging to an S-cell

The neocognitron is a network consisting of neuron-like cells of analog type; that is, their inputs and outputs take non-negative analog values corresponding to the instantaneous firing frequencies of biological neurons. Figure 6 shows the input-to-output characteristics of an S-cell, in which an inhibitory input reduces the effect of the excitatory inputs in a shunting manner.

In Figure 7, only the connections converging to an S-cell are shown (see also Figure 13 in the Appendix). The S-cell receives variable excitatory connections from a group of C-cells of the preceding layer. The cell also receives a variable inhibitory connection from an inhibitory cell, called a V-cell. The V-cell receives fixed excitatory connections from the same group of C-cells as does the S-cell, and always responds with the average intensity of the output of the C-cells.

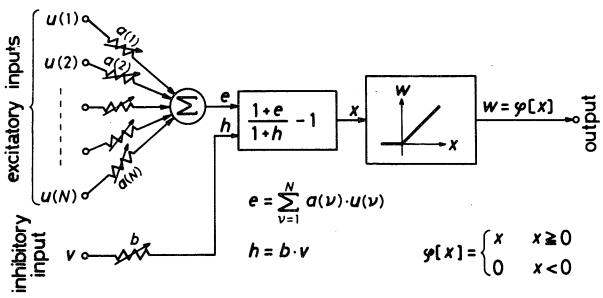

FIGURE 6. Input-to-output characteristics of an S-cell: A typical example of the cells employed in the neocognitron (Fukushima & Miyake, 1982).

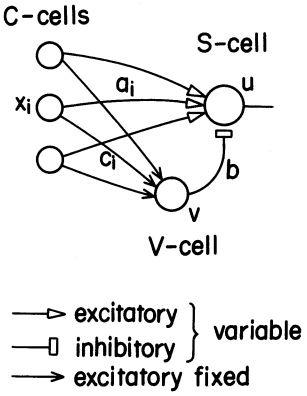

FIGURE 7. Connections converging to a feature-extracting S-cell.

Let the output of the S-cell be u. a_i is the strength of the connection from the ith C-cell, whose output is x_i. The S-cell also receives a variable inhibitory connection b from a V-cell, whose output is v. The inhibition works on the S-cell in a shunting manner.

Mathematically, the output of the S-cell is given by

$$u = r\varphi\left[\frac{1 + \sum_i a_i x_i}{1 + \frac{r}{1+r} b v} - 1\right] \tag{1}$$

where r is a positive constant determining the efficiency of the inhibition by the V-cell, and $\varphi[\ ]$ is a function defined by

$$\varphi[x] = \begin{cases} x & \text{if } x \geqq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{2}$$

The V-cell receives fixed excitatory connections c_i from the same group of C-cells as does the S-cell, and always responds with the average intensity (weighted root mean square) of the output of the C-cells:

$$v = \sqrt{\sum_i c_i \{x_i\}^2}. \tag{3}$$

If $bv \neq 0$, we can rewrite (1) as

$$u = \gamma\varphi\left[\frac{\sum_i a_i x_i}{b v} - \frac{r}{1+r}\right] \tag{4}$$

where γ is a function defined by

$$\gamma = (1 + r)\frac{\frac{r}{1+r} bv}{1 + \frac{r}{1+r} bv}. \tag{5}$$

In normal cases, in which the inhibitory connection b has already been reinforced to a large value, the value of γ is nearly equal to $(1 + r)$, which is a constant. That is,

$$\gamma \approx (1 + r) \quad \text{if } \frac{r}{1+r} bv \gg 1. \tag{6}$$

3.2. Reinforcement of Variable Connections

The neocognitron can be trained to recognize patterns either by unsupervised learning or by supervised learning. In this paper, we discuss only the case of unsupervised learning or learning-*without*-a-teacher.

Self-organization of the network is performed using the following principle: among the cells situated in a certain small area, only the one responding most strongly has its input connections reinforced. The amount of reinforcement of each input connection to this maximum-output cell is proportional to the intensity of the response of the cell from which the relevant connection leads. According to recent terminology, this principle can be classified under the competitive-learning paradigm.

Figure 8 illustrates this process of reinforcement, showing only the connections converging to an S-cell. The initial strength of these variable connections is nearly zero. Strictly speaking, each S-cell has very weak and diffused excitatory connections only during the initial period of self-organization. Once reinforcement of the input connections begins, the weak and diffused initial connections disappear.

Suppose the S-cell responds most strongly of the S-cells in its vicinity when a training stimulus is presented [see Figure 8(b)]. According to the principle described above, variable connections leading from activated C- and V-cells are reinforced, as shown in Figure 8(c). The variable excitatory connections to the S-cell grow into a "template" that exactly matches the spatial distribution of the response of the cells in the preceding layer. The inhibitory variable connection from the V-cell is also reinforced at the same

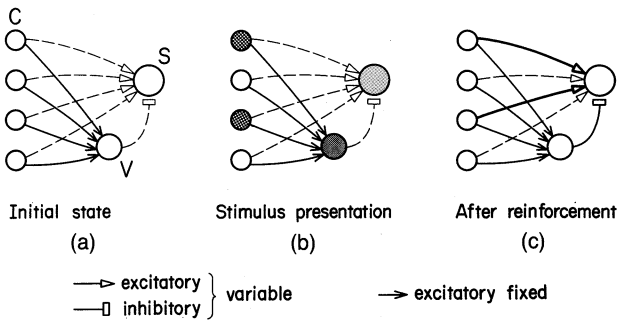

FIGURE 8. The process of reinforcement of the forward connections converging to a feature-extracting S-cell (Fukushima, 1988a). The density of the shadow in the circle represents the intensity of the response of the cell. (a) Shows the initial state before training; (b) shows stimulus presentation during the training; and (c) shows the connections after reinforcement.

time, but not strongly, because the output of the V-cell is not as large.

After training, the S-cell acquires the ability to extract a feature of the stimulus presented during the training period. Through the excitatory connections, the S-cell receives signals indicating the existence of the relevant feature to be extracted. If an irrelevant feature is presented, the inhibitory signal from the V-cell becomes stronger than the direct excitatory signals from the C-cells, and the response of the S-cell is suppressed. The S-cell is activated only when the relevant feature is presented. We could say that the V-cell "watches" for the existence of irrelevant features. Thus, inhibitory V-cells plays an important role in endowing the feature-extracting S-cells with the ability to differentiate irrelevant features, and in increasing the selectivity of feature extraction.

According to this principle, among the S-cells in a certain small area, only the one that yields a maximum output has its input connections reinforced. Because of the "winner-take-all" nature of this principle, the duplicate formation of cells extracting the same feature does not occur, and the formation of a redundant network can be prevented. only the one cell giving the greatest response to a training stimulus is selected, and only that cell is reinforced so as to respond more strongly to the stimulus.

Once a cell is selected and reinforced to respond to a feature, the cell usually loses its responsiveness to other features. When a different feature is presented, a different cell usually yields the maximum output and has its input connections reinforced. Thus, a "division of labor" among the cells occurs automatically.

Mathematically, the amount of reinforcement of each input connection to the maximum-output S-cell is proportional to the intensity of the response of the C-cell from which the relevant connection leads:

$$\Delta a_i = q\, c_i\, x_i \tag{7}$$

$$\Delta b = q\, v = q \sqrt{\sum_i c_i \{x_i\}^2} \tag{8}$$

where q is a positive constant determining the speed of reinforcement.

3.3. Analysis in Multidimensional Vector Space

We will now use vector notation, such as

$$\boldsymbol{x} = (x_1, x_2, x_3, \cdots)$$

$$\boldsymbol{a} = (a_1, a_2, a_3, \cdots)$$

$$\boldsymbol{c} = (c_1, c_2, c_3, \cdots)$$

to represent the response of the cells or the strength of the connections.

Here we define the weighted inner product of two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ by

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_i c_i\, x_i\, y_i \tag{9}$$

also using a weighting vector $\mathbf{c}$ whose elements are all positive ($c_i > 0$).

Using this inner product, we also define the norm (or the length) of a vector by

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}. \tag{10}$$

Let the S-cell have been reinforced by stimuli

$$\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(3)}, \cdots$$

and let

$$\boldsymbol{X} = \sum_n \boldsymbol{x}^{(n)} \quad \text{or} \quad X_i = \sum_n x_i^{(n)}. \tag{11}$$

Suppose that the initial strength of the variable connections $\boldsymbol{a}$ and b is zero, we have from (7) and (8)

$$a_i = q\, c_i \sum_n x_i^{(n)} = q\, c_i\, X_i \tag{12}$$

$$b = q \sqrt{\sum_n \sum_i c_i \{x_i^{(n)}\}^2} = q \sum_n \|\boldsymbol{x}^{(n)}\|. \tag{13}$$

Substituting (12) and (13) in (1), we obtain the next equation to represent the response of the S-cell:

$$u = \gamma\varphi \left[\lambda s - \frac{r}{1 + r} \right] \tag{14}$$

where s and λ are variables defined by

$$s = \frac{(\boldsymbol{X}, \boldsymbol{x})}{\|\boldsymbol{X}\| \cdot \|\boldsymbol{x}\|} \tag{15}$$

$$\lambda = \frac{\|\boldsymbol{X}\|}{\sum_n \|\boldsymbol{x}^{(n)}\|}. \tag{16}$$

The variable s represents a kind of similarity between the two vectors $\boldsymbol{x}$ and $\boldsymbol{X}$ in multidimensional vector space. After learning, λ becomes a positive constant independent of $\boldsymbol{x}$, and takes a value nearly equal to 1.0, as discussed later.

Since s defined by (15) is the inner product of the two vectors normalized by their norms, we can easily see that $s = 1$ if and only if the input vector $\boldsymbol{x}$ is in the same direction as $\boldsymbol{X}$, the sum of the training vectors. When $\boldsymbol{x}$ is not in the same direction as $\boldsymbol{X}$, we have $s < 1$.

We can see from (14) that we have $u > 0$ for an $\boldsymbol{x}$ whose similarity to $\boldsymbol{X}$ is great enough to satisfy

$$s > \frac{r}{1 + r} \cdot \frac{1}{\lambda}. \tag{17}$$

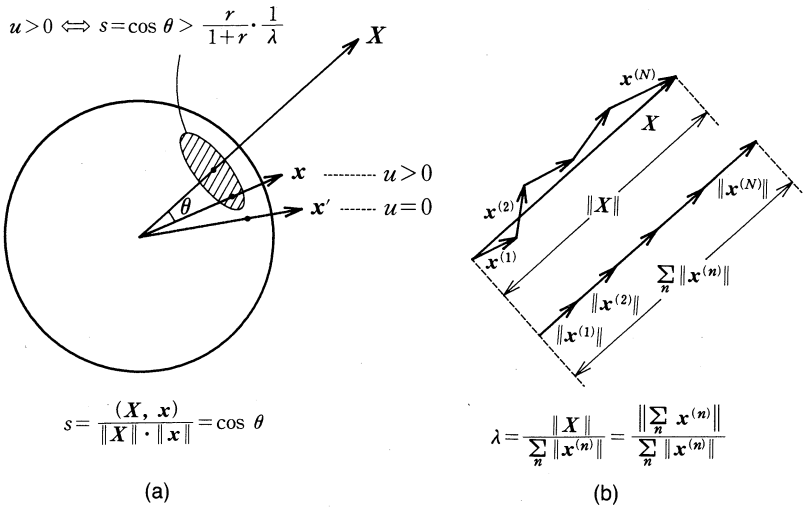

FIGURE 9. Similarity between patterns in a multidimensional vector space.

In other words, the S-cell yields a positive output if the direction of the input vector $\boldsymbol{x}$ falls within the shaded area in multidimensional vector space shown in Figure 9(a). However, if the input vector is directed outside this area, the S-cell does not respond.

The size of the shaded area is less for a larger value of the parameter r, so we can adjust the selectivity of the response of the S-cell by changing the value of the parameter r, and we can control the tolerance of distortion and noise. That is, a larger r gives a smaller tolerance of distortion and noise.

As can be seen from (16), the value of λ is in the range $\lambda \leqq 1$. Especially in the case of unsupervised learning, however, $\lambda \approx 1$ usually holds for the following reason. The S-cell learns the vectors which satisfy $u > 0$ only. Such a vector that yields $u = 0$ is not learned by the cell, because the cell cannot be the maximum output cell for that pattern. This means that the S-cell learns only training vectors in almost the same direction. Since all the training vectors $\boldsymbol{x}^{(n)}$ learned by the cell are in almost the same direction, the norm of the sum of $\boldsymbol{x}^{(n)}$ has almost the same value as the sum of the norms of $\boldsymbol{x}^{(n)}$, as shown in Figure 9(b). Hence their ratio λ is nearly equal to 1.

Although we usually have $\lambda < 1$, we can prove that the value of $r/\{(1 + r)\lambda\}$ on the right side of eqn (17) never becomes larger than 1, as long as the parameter r is kept constant during learning. If the value became larger than 1, the cell would lose its responsiveness and never respond to any later input. However this is not the case. A vector that would make the value larger than 1 elicits the response $u = 0$ from the cell, and cannot be learned. This means that in unsupervised learning of the neocognitron, it is always guaranteed that a cell never loses responsiveness.

4. DEVELOPMENT OF ITERATIVE CONNECTIONS:

Another important principle introduced for the self-organization of the neocognitron is that the maximum-output S-cell not only grows, but also controls the growth of neighboring cells, working, so to speak, like a seed in crystal growth. Neighboring S-cells have their input connections reinforced in the same way as the seed cell.

When a seed cell is selected from a cell-plane of S-cells, all the other S-cells in the cell-plane grow to have input connections of the same spatial distribution as the seed cell. As a result, all the S-cells in a cell-plane grow to receive input connections of the identical spatial distribution where only the positions of the preceding C-cells are shifted in parallel with the position of the S-cells. Because connections develop iteratively in a cell-plane, all the S-cells in the cell-plane come to respond selectively to a particular feature. Differences among these cells arise only from differences in the position of the feature to be extracted.

5. THE ROLE OF C-CELLS

It should be noted that, in the similarity criterion s defined by (15), spatial information about the position on the two-dimensional retina is completely ignored. For example, the similarity s between the two patterns in Figure 10, namely the pattern consisting of circles and the pattern consisting of crosses, is 0 according to this criterion, because they are orthogonal to each other in multidimensional vector space. They are quite similar to each other, however, when they are observed as two-dimensional spatial patterns.

The C-cells play an important role in handling a

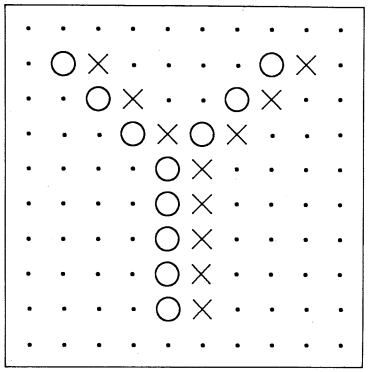

FIGURE 10. Two patterns, one of which consists of circles and the other of crosses, resemble each other closely if looked at as two-dimensional spatial patterns, but are orthogonal in multidimensional vector space.

stimulus as a two-dimensional spatial pattern. The C-cells are inserted in the network to allow for positional errors of the features extracted by the preceding S-cells.

Connections converging to C-cells are fixed and invariable. Each C-cell receives excitatory signals from a group of S-cells which extract the same feature, but from slightly different positions as shown in Figure 11(a). The C-cell is activated if at least one of these S-cells is active.

If we look at this phenomenon in a different way, we can see that the excitation of a single S-cell will elicit positive responses from a number of C-cells, as shown in Figure 11(b). This means that the output of the C-cells becomes a "blurred" version of the S-cell's output.

Suppose the input pattern is slightly deformed or shifted from the training pattern. Usually, the response to the input pattern and the response to the training pattern do not overlap in the layer of S-cells. This means that the response vectors to these patterns are orthogonal to each other in multidimensional vector space.

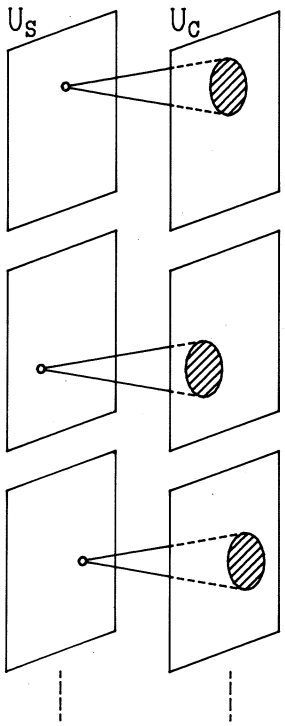

FIGURE 12. "Blurring" operation performed at each cell-plane in a layer of C-cells.

In the layer of C-cells, the response of the layer of S-cells is "blurred." Hence, the response to the input and to the training patterns usually overlap in the layer of C-cells. This means that the similarity

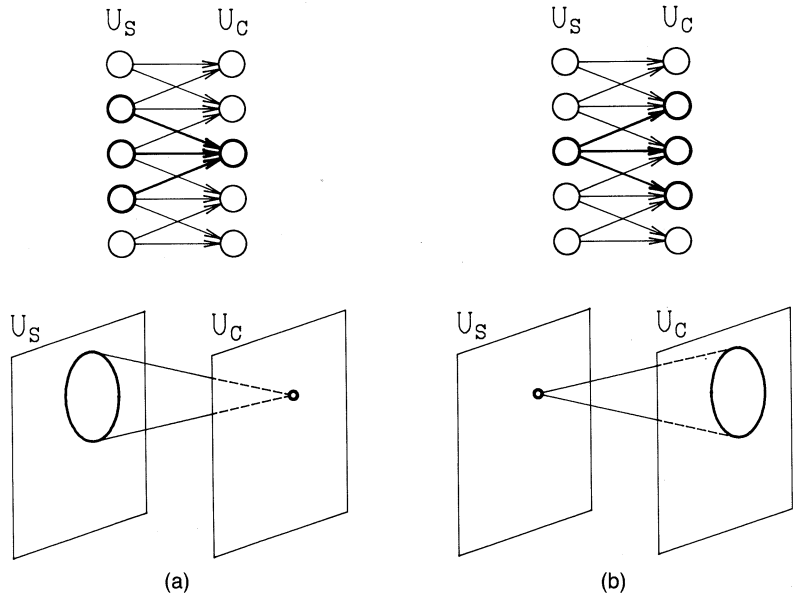

FIGURE 11. Connections from S-cells to C-cells.

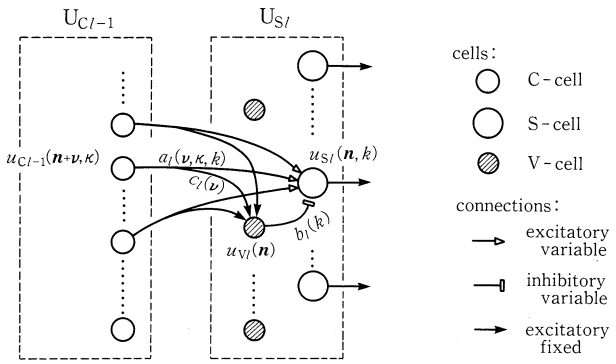

FIGURE 13. Connections between two adjoining layers of C-cells and S-cells.

defined by s becomes large enough to elicit a positive response from an S-cell in the next layer.

For instance, the similarity between the two patterns shown in Figure 10 can be made large by a "blurring" operation. If the "blurring" operation is applied directly to the input patterns, however, the detailed structure of the input patterns is also blurred, and the difference in such detailed structures cannot be distinguished. Hence in the neocognitron, the "blurring" operation is performed by C-cells after local features have been extracted by S-cells. The "blurring" operation is performed independently in each cell-plane as shown in Figure 12.

If we refer to the transfer characteristics of the output of an S-cell from the output of the S-cells in the preceding layer via the intermediate layer of C-cells, we can say roughly that the S-cell judges the similarity between the stimulus and the training patterns by observing the degree of overlap between the blurred versions.

6. NECESSARY NUMBER OF CELLS IN THE NETWORK

The number of cells (or to be more exact, the number of cell-planes) in each stage of the network must be increased when the number of categories of the patterns to be recognized is increased.

However, the increase is less than linear because local features to be extracted at the lower stages are usually shared by several patterns of different categories.

REFERENCES

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

Fukushima, K. (1988a). A neural network for visual pattern recognition. *IEEE Computer*, **21**(3), 65–75.

Fukushima, K. (1988b). Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, **1**(2), 119–130.

Fukushima, K., & Miyake, S. (1982). Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, **15**(6), 455–469.

APPENDIX: MATHEMATICAL DESCRIPTION OF THE WHOLE NETWORK

The initial stage of the network is the input layer called U_0, and consists of a two-dimensional array of receptor cells u_0. The layers of S-cells and C-cells at the lth stage are denoted by U_{Sl} and U_{Cl}, respectively. The notation $u_{Sl}(\boldsymbol{n}, k)$, for example, is used to denote the output of an S-cell in layer U_{Sl}, where $\boldsymbol{n}$ is a two-dimensional set of coordinates indicating the position of the cell's receptive-field center in the input layer U_0, and k is the serial number of the cell-plane ($1 \leqq k \leqq K_l$).

Figure 13 illustrates the connections to a layer of S-cells from the layer of C-cells at the preceding stage. The output of an S-cell is given by

$$u_{Sl}(\boldsymbol{n}, k) = r_l \times \varphi\left[\frac{1 + \sum_{\kappa=1}^{K_{l-1}} \sum_{\nu \in A_l} a_l(\nu, \kappa, k) \cdot u_{Cl-1}(\boldsymbol{n} + \nu, \kappa)}{1 + \frac{r_l}{1 + r_l} \cdot b_l(k) \cdot u_{Vl}(\boldsymbol{n})} - 1\right] \quad \text{(A.1)}$$

where $\varphi[\ ]$ is the function defined by (2). In the case of $l = 1$ in (1), $u_{Cl-1}(\boldsymbol{n}, \kappa)$ stands for $u_0(\boldsymbol{n})$ or the output of a receptor cell of the input layer, and we have $K_{l-1} = 1$.

$a_l(\nu, \kappa, k)$ ($\geqq 0$) is the strength of the variable excitatory connection coming from C-cell $u_{Cl-1}(\boldsymbol{n} + \nu, \kappa)$ of the preceding stage. A_l denotes the summation range of ν, that is, the size of the spatial spread of the input connections to one S-cell. $b_l(k)$ ($\geqq 0$) is the strength of the variable inhibitory connection coming from subsidiary V-cell $u_{Vl}(\boldsymbol{n})$. As discussed above in sections 2 and 4, all the S-cells in a cell-plane have identical sets of input connections. Hence, $a_l(\nu, \kappa, k)$ and $b_l(k)$ do not contain argument $\mathbf{n}$ representing the position of the receptive field of the cell $u_{Sl}(\boldsymbol{n}, k)$. The positive constant r_l determines the efficiency of the inhibitory input to this cell.

The subsidiary V-cell which sends an inhibitory signal to this S-cell yields an output equal to the weighted root-mean-square of the signals from the preceding C-cells; that is,

$$u_{Vl}(\boldsymbol{n}) = \sqrt{\sum_{\kappa=1}^{K_{l-1}} \sum_{\nu \in Al} c_l(\nu) \cdot \{u_{Cl-1}(\boldsymbol{n} + \nu, \kappa)\}^2}, \quad \text{(A.2)}$$

where $c_l(\nu)$ represents the strength of the fixed excitatory connections, and is a monotonically decreasing function of $\|\nu\|$.

The output of a C-cell inserted in the network to allow for positional errors, is given by

$$u_{Cl}(\boldsymbol{n}, k) = \psi\left[\sum_{\nu \in D_l} d_l(\nu) \cdot u_{Sl}(\boldsymbol{n} + \nu, k)\right], \quad \text{(A.3)}$$

where $\psi[\ ]$ is a function specifying the characteristic of saturation of the C-cell, and is defined by

$$\psi[x] = \frac{\varphi[x]}{1 + \varphi[x]}. \quad \text{(A.4)}$$

Parameter $d_l(\nu)$ denotes the strength of the fixed excitatory connections, and is a monotonically decreasing function of $\|\nu\|$. D_l is the area to which these connections spread.

During learning, the variable connections $a_l(\nu, \kappa, k)$ and $b_l(k)$ are reinforced depending on the intensity of the input to the seed cell. Let $u_{Sl}(\hat{\boldsymbol{n}}, \hat{k})$ be selected as a seed cell at a certain time. The variable connections $a_l(\nu, \kappa, \hat{k})$ and $b_l(\hat{k})$ to this seed cell, and consequently to all the S-cells in the same cell-plane as the seed cell, are reinforced by the following amount:

$$\Delta a_l(\nu, \kappa, \hat{k}) = q_l \cdot c_l(\nu) \cdot u_{Cl-1}(\hat{\boldsymbol{n}} + \nu, \kappa), \quad \text{(A.5)}$$

$$\Delta b_l(\hat{k}) = q_l \cdot u_{Vl}(\hat{\boldsymbol{n}}), \quad \text{(A.6)}$$

where q_l is a positive constant determining the speed of reinforcement.