Pergamon

Neural Networks, Vol 7, No 1, pp 41–47, 1994
Copyright © 1994 Elsevier Science Ltd
Printed in the USA All rights reserved
0893-6080/94 $6 00 + 00

CONTRIBUTED ARTICLE

Neocognitron With Dual C-Cell Layers

KUNIHIKO FUKUSHIMA, MASATO OKADA, AND KAZUHITO HIROSHIGE

Osaka University

(*Received* 5 *October* 1992; *revised and accepted* 22 *April* 1993)

Abstract—*The neocognitron is a hierarchical neural network model capable of deformation-resistant pattern recognition In the hierarchical network of the neocognitron, feature extraction by the* S-*cells and a blurring operation by the* C-*cells are repeated. The ability of each* S-*cell to robustly extract deformed features is created by the blurring operation of the* C-*cells placed in front of the* S-*cell*

In the conventional neocognitron, the amount of blurring produced by the C-*cells is uniform in the receptive field of each* S-*cell. An* S-*cell would accept a much larger deformation if a nonuniform blurring could be produced in such a way that a larger blurring is generated in the periphery than at the center of the receptive field This is desirable as discrepancies between a training pattern and a deformed stimulus pattern usually become larger in the periphery than at the center of the receptive field*

In order to produce such a nonuniform blurring economically, we propose a neocognitron with a dual C-*cell layer A layer of* C-*cells in an intermediate stage of the network is divided into two sub-layers one with a smaller blurring and the other with a larger blurring Each* S-*cell in the succeeding layer receives input connections from the low-blur* C-*cell layer at the center of its connecting area while also receiving connections from the high-blur* C-*cell layer at its periphery Computer simulation has shown that the new neocognitron recognizes characters more robustly than the conventional neocognitron.*

Keywords—Neocognitron, Neural network, Visual pattern recognition, Character recognition, Dual C-cell layer, Nonuniform blurring.

1. INTRODUCTION

The neocognitron is a hierarchical neural network model that is capable of deformation-resistant pattern recognition (Fukushima, 1980, 1988). It can acquire an extremely robust pattern recognition ability through the process of learning. Because its generalization ability is large, the neocognitron requires relatively short training time compared to other artificial neural networks trained by backpropagation (e.g., LeCun, Boser, Denker, et al., 1989).

In the hierarchical network of the neocognitron, feature extraction by the S-cells and a blurring operation by the C-cells are repeated. The ability of each S-cell to robustly extract deformed features is created by the blurring operation of the C-cells placed in front of the S-cell (Fukushima, 1989). This is due to the fact that positional errors of the local features extracted by the S-cells in the preceding stage can be tolerated as a result of the blurring operation of the C-cells.

Consider an S-cell that has previously been trained to extract a feature such as a corner or a crossing of two lines. If a deformed version of the training pattern is presented as a stimulus to the input layer, the discrepancy between the feature in the training pattern and that of the stimulus pattern usually becomes larger in the periphery than at the center of the receptive field of the S-cell. In the conventional neocognitron, however, the amount of blurring produced by the C-cells is uniform throughout the receptive field of the S-cell. It can be expected that the S-cell might accept much larger deformation of the feature, if a nonuniform blurring could be produced in such a way that a larger blurring is generated in the periphery than at the center of the receptive field.

A straightforward way of realizing this nonuniform blurring operation is to prepare for each individual S-cell a set of C-cells of its own, and to make the C-cells at the center have a smaller blurring while the C-cells near the periphery have a larger blurring depending on the distance from the center. This is not practical, however, because of the enormous number of C-cells required by the network.

Acknowledgements: This work was supported in part by Grant-in-Aid #02402035 for Scientific Research (A), and #04246105 for Scientific Research on Priority Areas on "Higher-Order Brain Functions," both from the Ministry of Education, Science and Culture of Japan.

Requests for reprints should be sent to Prof. Kunihiko Fukushima, Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University, Osaka, Japan.

In order to get a similar effect much more economically, we propose a neocognitron with dual C-cell layers (Hiroshige, Okada, & Fukushima, 1992). Each layer of C-cells is divided into two sublayers: one with a smaller blurring and the other with a larger blurring. Each S-cell in the succeeding layer is constructed to receive input connections from the low-blur C-cell layer at the center of its connecting area, and also receive connections from the high-blur C-cell layer in the periphery of its connecting area.

The neocognitron with this new architecture has been simulated on a computer, and has been trained to recognize handwritten numeric characters. It is shown that the new neocognitron recognizes characters more robustly than the conventional neocognitron.

2. CONVENTIONAL NEOCOGNITRON

Because the network proposed in this paper is a modified version of the conventional neocognitron, the network architecture of the conventional neocognitron (Fukushima, 1988) is briefly described in this section. Readers who are familiar with the conventional neocognitron and the notation in its equations may skip this section.

Figure 1 shows the network architecture of the conventional neocognitron. The lowest stage of the network is the input layer U_0. Each of the succeeding stages has a layer U_S consisting of "S-cells" followed by another layer U_C of "C-cells." The layers U_S and U_C at the lth stage are denoted by U_{Sl} and U_{Cl}, respectively.

Each layer of S-cells or C-cells is divided into subgroups, called "cell-planes," according to the features to which they respond. Each rectangle drawn with heavy lines in Figure 1 represents a cell-plane. The connections converging to the cells in a cell-plane are homogeneous and topographically ordered. In other words, the connections have a translational symmetry. (We can also say that all the cells of a cell-plane share the same set of input connections.) Each cell receives its input connections from only a limited number of cells situated in a small area on the preceding layer.

The density of cells in each cell-plane is designed to decrease with the order of the stage. Layer U_{C4} at the highest stage is the recognition layer, representing the

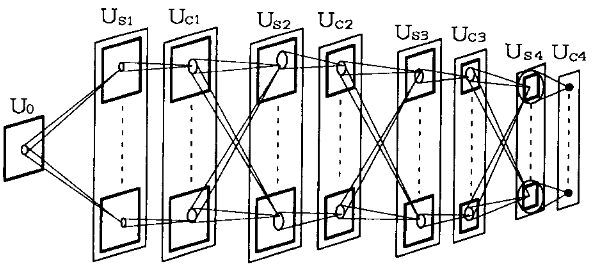


FIGURE 1. The hierarchical network architecture of the conventional neocognitron.

final result of the pattern recognition. In this layer, each cell-plane contains only one C-cell.

The optimal scale of the network changes depending on the set of patterns to be recognized (Fukushima & Wake, 1991). If the complexity of the patterns is high, the total number of stages in the network needs to be large.

In the mathematical descriptions below, 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 ($1 \le k \le K_l$).

The S-cells are feature-extracting cells. Each layer U_{Sl} contains subsidiary V-cells, as well as S-cells. The V-cells send inhibitory signals to the S-cells. The S-cells, with the aid of the V-cells, extract features from the input pattern. Mathematically, the outputs of an S-cell and the subsidiary V-cell are given by

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

$$u_{Vl}(\boldsymbol{n}) = \left\{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{\boldsymbol{\nu}\in A_l} c_l(\boldsymbol{\nu}) \cdot [u_{Cl-1}(\boldsymbol{n}+\boldsymbol{\nu}, \kappa)]^2\right\}^{1/2}, \quad (2)$$

where $\varphi[\ ]$ is a function defined by $\varphi[x] = \max(x, 0)$. In the case of $l = 1$ in eqn (1), $u_{Cl-1}(\boldsymbol{n}, \kappa)$ stands for $u_0(\boldsymbol{n})$ or the output of a receptor cell of the input layer U_0, and we have $K_{Cl-1} = 1$.

Parameter $a_l(\boldsymbol{\nu}, \kappa, k)$ (≥ 0) is the strength of the variable excitatory connection coming from C-cell $u_{Cl-1}(\boldsymbol{n} + \boldsymbol{\nu}, \kappa)$ of the preceding stage. A_l denotes the summation range of $\boldsymbol{\nu}$, that is, the size of the spatial spread of the input connections to one S-cell. Parameter $b_l(k)$ (≥ 0) is the strength of the variable inhibitory connection coming from the V-cell. Because all the S-cells in a cell-plane have identical sets of input connections, $a_l(\boldsymbol{\nu}, \kappa, k)$ and $b_l(k)$ do not contain argument $\boldsymbol{n}$ 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 S-cell. Parameter $c_l(\boldsymbol{\nu})$ represents the strength of the fixed excitatory connections to the V-cell, and is a monotonically decreasing function of $|\boldsymbol{\nu}|$.

C-cells are put into the network to allow for positional error in the features extracted by the S-cells. The connections from S-cells to C-cells are fixed. Each C-cell receives signals from a group of S-cells that extract the same feature, but from slightly different positions. It is activated if at least one of these S-cells is active. Mathematically, the response of a C-cell of U_{Cl} is given by

$$u_{Cl}(\boldsymbol{n}, k) = \psi\left[\sum_{\kappa=1}^{K_{Sl}} j_l(\kappa, k) \sum_{\boldsymbol{\nu}\in D_l} d_l(\boldsymbol{\nu}) \cdot u_{Sl}(\boldsymbol{n}+\boldsymbol{\nu}, \kappa)\right], \quad (3)$$

where $\psi[x] = \varphi[x]/(1 + \varphi[x])$. Parameter $d_l(\nu)$ represents the strength of the fixed excitatory connections, which spread within D_l, and is a monotonically decreasing function of $|\nu|$.

In the network trained by supervised learning, an operation called "join" is performed (Fukushima, 1988; Fukushima & Wake, 1991). If the deformation of a feature is too large to be extracted by a single S-cell-plane, another S-cell-plane is created to extract a deformed version of the same feature. The output of the second S-cell-plane is joined to the output of the first S-cell-plane and they are fed to a single C-cell-plane together. Parameter $j_l(\kappa, k)$ represents how the S-cell-planes are joined. It takes the value 1 if the κth cell-plane is to be joined to the kth cell-plane, and takes 0 otherwise. In the network trained by unsupervised learning, in which no joining operation is performed, we have $j_l(\kappa, k) = 1$ for $\kappa = k$, and $j_l(\kappa, k) = 0$ for $\kappa \neq k$.

During the training, 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. Seed cells are appointed by the "teacher" in the case of supervised learning (Fukushima, 1988; Fukushima & Wake, 1991), or determined by a kind of winner-take-all process in the case of unsupervised learning (Fukushima & Miyake, 1982). If cell $u_{Sl}(\hat{\boldsymbol{n}}, \hat{k})$ is 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:

$$\Delta a_l(\nu, \kappa, \hat{k}) = q_l \cdot c_l(\nu) \cdot u_{Cl-1}(\hat{\boldsymbol{n}} + \nu, \kappa), \quad (4)$$

$$\Delta b_l(\hat{k}) = q_l \cdot u_{Vl}(\hat{\boldsymbol{n}}), \quad (5)$$

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

3. EFFECT OF NONUNIFORM BLURRING

Let us consider a case in which a neocognitron has already finished learning a character of a particular style of writing, for example, the pattern shown in Figure 2(a). Generally in an intermediate stage of the network, an S-cell has been created to extract, for example, the feature enclosed by the circle in Figure 2(a). It is desirable if the same S-cell can also respond to the corresponding feature in a deformed version of the same character, which is shown in the circle in Figure 2(b).

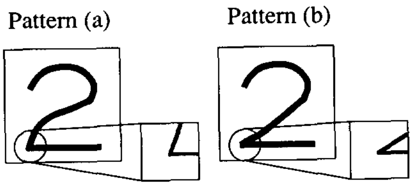


FIGURE 2. A training character and a deformed version of the same character.

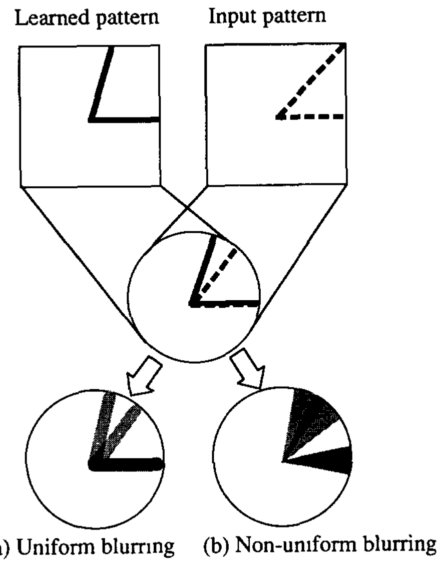


FIGURE 3. Effects of uniform and nonuniform blurring on the recognition of a deformed feature.

An S-cell usually judges the similarity between the learned and the input features by the degree of overlap of the blurred versions of the two features (Fukushima, 1989). If the blurring is uniform within the receptive field of the S-cell, there is not enough overlap between the two features as shown in Figure 3(a), and the S-cell will fail to respond to the deformed feature. This is because the discrepancy between the two features increases in the periphery of the receptive field. If we can create a larger blurring in the periphery of the receptive field, however, we can have enough overlap even for a deformed feature as shown in Figure 3(b).

Of course we can have a better overlap between the two features if we have high blurring throughout the receptive field. However, the S-cell then loses the ability to discriminate different but similar features, as shown in Figure 4. Therefore, it is essential to have low blurring near the center of the receptive field.

4. DUAL C-CELL LAYERS

A straightforward way of realizing this nonuniform blurring operation is to give each individual S-cell a set of C-cells of its own, and to make the C-cells at the center have a smaller blurring while the C-cells near the periphery have a larger blurring depending on the distance from the center. This is not practical, however, because of the enormous number of C-cells that would be required by this network.

In order to get the same effect much more economically, we propose a neocognitron with dual C-cell layers. Each layer of C-cells is divided into two sublayers: one with a smaller blurring and the other with a larger blurring. Each S-cell in the succeeding layer is constructed to receive input connections from the low-blur C-cell layer at the center of its connecting area, and

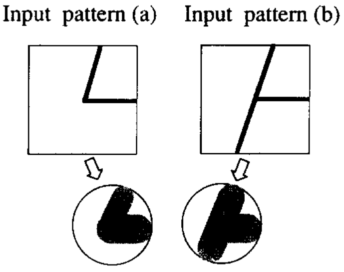


FIGURE 4. Impairment of the ability to discriminate different features of similar shape when blurring is too large throughout the receptive field.

also receive connections from the high-blur C-cell layer in the periphery of its connecting area.

We first applied this operation to layer U_{C1} only, as discussed in the next section. Figure 5 illustrates the architecture of the network to realize this operation. Layer U_{C1} is divided into two sublayers, U^{α}_{C1} and U^{β}_{C1}.

Figure 6 shows the one-dimensional cross section of the connections to and from layers U^{α}_{Cl} and U^{β}_{Cl} ($l = 1$ in this case). Each C-cell of U^{α}_{Cl} receives input connections from a smaller number of S-cells of U_{Sl} and has a smaller blurring than the C-cells of U^{β}_{Cl}. The spatial distributions of the input connections of these two sets of C-cells are represented by $d^{\alpha}_{l}(\nu)$ and $d^{\beta}_{l}(\nu)$, and the size of the spatial spread of these connections by D^{α}_{l} and D^{β}_{l}. As can be seen from the figure, we have $D^{\alpha}_{l} \subset D^{\beta}_{l}$.

Mathematically, the response of C-cells of U^{α}_{Cl} and U^{β}_{Cl} are given by

$$u^{\alpha}_{Cl}(\mathbf{n},k) = \psi\left[\sum_{\kappa=1}^{K_{Sl}} j_l(\kappa,k) \sum_{\nu \in D^{\alpha}_{l}} d^{\alpha}_{l}(\nu)\cdot u_{Sl}(\mathbf{n}+\nu,\kappa)\right], \tag{6}$$

$$u^{\beta}_{Cl}(\mathbf{n},k) = \psi\left[\sum_{\kappa=1}^{K_{Sl}} j_l(\kappa,k) \sum_{\nu \in D^{\beta}_{l}} d^{\beta}_{l}(\nu)\cdot u_{Sl}(\mathbf{n}+\nu,\kappa)\right] \tag{7}$$

Each S-cell of the succeeding layer U_{Sl+1} receives connections from C-cells of layer U^{α}_{Cl} and U^{β}_{Cl}. At the center of its connecting area, it receives stronger connections from layer U^{α}_{Cl} than layer U^{β}_{Cl}; while in the periphery, it receives stronger connections from layer U^{β}_{Cl} than layer U^{α}_{Cl}. In order to realize this condition, the weighting functions $c^{\alpha}_{l+1}(\nu)$ and $c^{\beta}_{l+1}(\nu)$, which determine the degree of effect from an individual point in the connecting area, are so determined as to have the spatial distributions shown in Figure 6.

Mathematically, the outputs of an S-cell and the subsidiary V-cell that sends the inhibitory signal to this S-cell, are given by

$$u_{Sl+1}(\mathbf{n},k) = r_{l+1}\cdot\varphi\left[\frac{1+\iota_{Sl+1}(\mathbf{n},k)}{1+\dfrac{r_{l+1}}{1+r_{l+1}}\cdot b_{l+1}(k)\cdot u_{Vl+1}(\mathbf{n})} - 1\right], \tag{8}$$

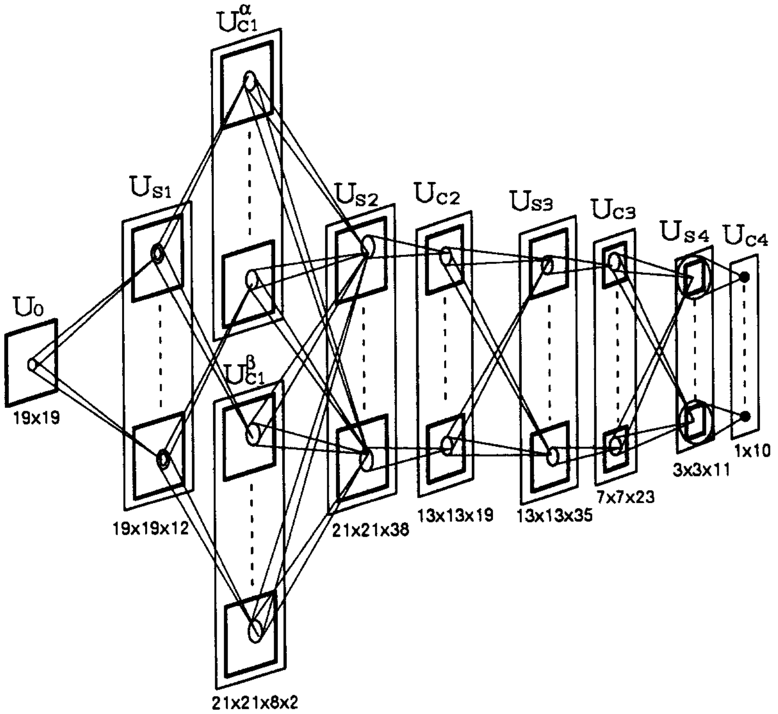


FIGURE 5. The architecture of the network containing a dual C-cell layer. The numerals at the bottom of the figure show the total number of S- or C-cells (not including V-cells) in the individual layers of the network. The number of V-cells in each layer is the same as that of the S-cells in one cell-plane of that layer. For example, U_{S1} has 19 × 19 V-cells.

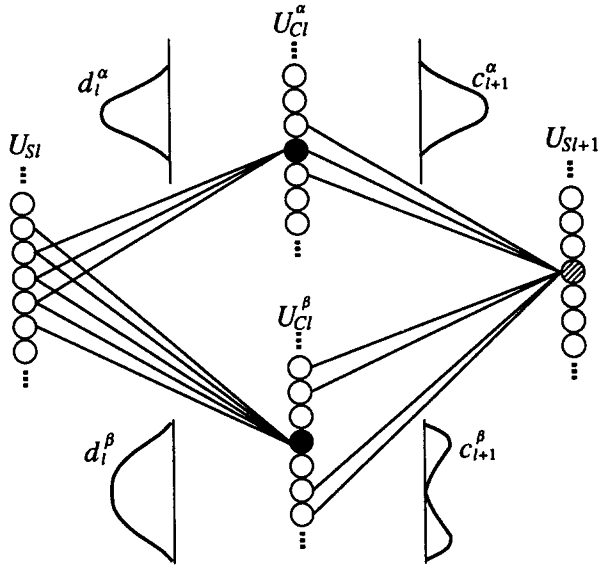


FIGURE 6. One-dimensional cross section of the connections to and from layers U_{Cl}^{α} and U_{Cl}^{β}.

where

$$i_{Sl+1}(\mathbf{n}, k) = \sum_{\kappa=1}^{K_{Cl}} \sum_{\nu \in A_{l+1}} [a_{l+1}^{\alpha}(\nu, \kappa, k) \cdot u_{Cl}^{\alpha}(\mathbf{n} + \nu, \kappa) + a_{l+1}^{\beta}(\nu, \kappa, k) \cdot u_{Cl}^{\beta}(\mathbf{n} + \nu, \kappa)], \quad (9)$$

and

$$u_{Vl+1}(\mathbf{n}) = \left\{ \sum_{\kappa=1}^{K_{Cl}} \sum_{\nu \in A_{l+1}} c_{l+1}^{\alpha}(\nu) \cdot [u_{Cl}^{\alpha}(\mathbf{n} + \nu, \kappa)]^2 + c_{l+1}^{\beta}(\nu) \cdot [u_{Cl}^{\beta}(\mathbf{n} + \nu, \kappa)]^2 \right\}^{1/2}. \quad (10)$$

During training, the variable connections $a_{l+1}^{\alpha}(\nu, \kappa, k)$, $a_{l+1}^{\beta}(\nu, \kappa, k)$, and $b_{l+1}(k)$ are reinforced depending on the intensity of the input to the seed cell. If cell $u_{Sl+1}(\hat{\mathbf{n}}, \hat{k})$ is selected as a seed cell at a certain time, the variable connections $a_{l+1}^{\alpha}(\nu, \kappa, \hat{k})$, $a_{l+1}^{\beta}(\nu, \kappa, \hat{k})$, and $b_{l+1}(\hat{k})$ 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:

$$\Delta a_{l+1}^{\alpha}(\nu, \kappa, \hat{k}) = q_{l+1} \cdot c_{l+1}^{\alpha}(\nu) \cdot u_{Cl}^{\alpha}(\hat{\mathbf{n}} + \nu, \kappa), \quad (11)$$

$$\Delta a_{l+1}^{\beta}(\nu, \kappa, \hat{k}) = q_{l+1} \cdot c_{l+1}^{\beta}(\nu) \cdot u_{Cl}^{\beta}(\hat{\mathbf{n}} + \nu, \kappa), \quad (12)$$

$$\Delta b_{l+1}(\hat{k}) = q_{l+1} \cdot u_{Vl+1}(\hat{\mathbf{n}}). \quad (13)$$

5. COMPUTER SIMULATION

5.1. Dualization of Layer U_{C1}

In order to test the effect of the dual C-cell layers, we first tried to dualize the C-cell layer of only the first stage, as shown in Figure 5.

The shape of the functions $d_1^{\alpha}(\nu)$ and $d_1^{\beta}(\nu)$, and the size of their spatial spread D_1^{α} and D_1^{β} in eqns (6) and (7), which specify the input connections of the C-cells, have been determined by the following equations:

$$d_1^{\alpha}(\nu) = \delta_{\alpha}^{|\nu|^2}, \quad D_1^{\alpha}: 3 \times 3, \quad (14)$$

$$d_1^{\beta}(\nu) = \delta_{\beta}^{|\nu|^2}, \quad D_1^{\beta}: 5 \times 5, \quad (15)$$

where δ_{α} and δ_{β} are positive constants ≤ 1. Incidentally, in the conventional neocognitron (Fukushima, 1988), we have

$$d_1(\nu) = 0.9^{|\nu|^2}, \quad D_1: 3 \times 3. \quad (16)$$

The shape of the weighting functions $c_2^{\alpha}(\nu)$ and $c_2^{\beta}(\nu)$ in eqns (10), (11), and (12), which specify the input connections of the S-cells, has been determined by the following equations:

$$c_2^{\alpha}(\nu) = \rho^{|\nu|^2}, \quad (17)$$

$$c_2^{\beta}(\nu) = c_2(\nu) - c_2^{\alpha}(\nu), \quad (18)$$

where $c_2(\nu)$ is the weighting function used in eqns (2) and (4) in the conventional neocognitron (Fukushima, 1988), and is defined by

$$c_2(\nu) = 0.9^{|\nu|^2}, \quad (19)$$

and ρ takes a value within the range of $0 < \rho \leq 0.9$. If we choose $\rho = 0.9$, we have $c_2^{\beta}(\nu) = 0$, and the S-cells of U_{S2} do not receive any signal from U_{C1}^{β}. In this case, the network reduces to the conventional neocognitron. With a smaller value of ρ, however, each S-cell comes to receive connections mainly from U_{C1}^{β} in the periphery of its receptive field.

In order to determine the optimum values of the parameters, we measured the performance of the network for various values of δ_{α}, δ_{β}, and ρ. The network was trained to recognize numeric characters from 0 to 9 by supervised learning using the same set of training patterns as the conventional neocognitron (Fukushima, 1988). After the learning process had been completed, the recognition error was measured.

In order to measure the recognition rate, we gathered 1000 samples of numeric characters handwritten by many subjects using different styles of writing. This set of patterns was randomly divided into two groups: one was to be used to find out the optimum values of the parameters, and the other to test if the chosen parameters were really optimal even for unknown patterns.

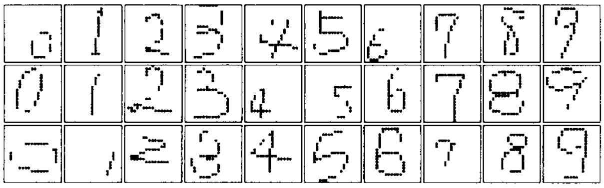
After the learning process was completed for each combination of different values of the parameters, the recognition error was measured using a set of 400 randomly chosen test patterns, which contained 40 deformed patterns for each character from 0 to 9. It was discovered that the error rate was minimized at $\delta_{\alpha} = 0.7$, $\delta_{\beta} = 0.9$, and $\rho = 0.8$. The error rate in recognizing the 400 test patterns with this set of parameters was

8.3%, while the error rate of the conventional neocognitron, which is obtained by setting $\delta_\alpha = 0.9$ and $\rho = 0.9$, was 17.8%. (In order to confirm that the parameters in the conventional neocognitron, to which our new network is to be compared, have been optimally chosen, we also tested the performance of the conventional neocognitron with different values of δ_α at $\rho = 0.9$, and found that the $\delta_\alpha = 0.9$ gives the best performance.)

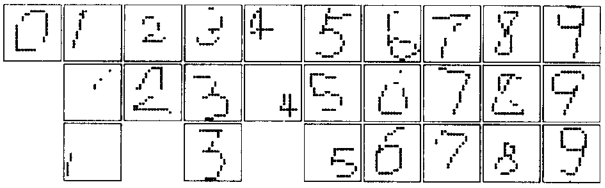
In order to confirm that the chosen parameters were really optimal even for unknown patterns, we measured the performance of the network using the 600 test patterns that had not been used in the previous process of determining the optimum parameter values. The error rate of the network using the optimum parameter values was 9.2%, while the error rate of the conventional neocognitron was 16.3%. These data show that the parameters have been chosen optimally even for unknown patterns.

Because the test patterns consist of characters of bad handwriting, the absolute values of these recognition rates do not have any significance. Only the relative value is important here, for purposes of comparison. We can see that the error rate can be halved by the dualization of the C-cell layer. From these results, we can conclude that the dualization of layer U_{C1} is effective.

Figure 7 shows some examples of deformed numeric characters that the network with optimal values of δ_α, δ_β, and ρ has recognized correctly. These patterns have been chosen from the set of 600 unknown patterns. Among these patterns, the patterns shown in (a) were recognized correctly both by the new network and the conventional neocognitron, while the patterns in (b) were recognized correctly only by the new network and not by the conventional neocognitron.



(a)



(b)

FIGURE 7. Some examples of deformed numeric characters that the network recognized correctly. (a) Patterns recognized correctly both by the new network and the conventional neocognitron. (b) Patterns recognized correctly only by the new network.

5.2. Dualization of Other C-Cell Layers

Because the error rate has been reduced by the dualization of layer U_{C1}, we can also expect to have a better performance by dualizing other layers of C-cells. Therefore, we attempted to dualize layer U_{C2}, as well as layer U_{C1}, and measured the error rate by changing the value of ρ for U_{C2}. However, no distinct improvements in the recognition rate by dualizing U_{C2} were realized. The reason why dualization of the C-cell layer is not as effective in the higher stages may be as follows.

Each S-cell of U_{S3} is usually trained to extract a global feature consisting of a combination of local features, and the positional errors of these local features are absorbed by the blurring operation of the C-cells of U_{C2}. These local features are not necessarily placed at the center of the receptive field of the S-cell of U_{S3}, but are generally placed in the periphery. Hence, a reduced blurring at the center is not as effective as for the S-cells of U_{S2}, for which important local features are usually placed at the center.

6. DISCUSSION

We have dualized a layer of C-cells of the neocognitron, and produced a nonuniform blurring of the input to S-cells. With this modification, S-cells are able to accept a much larger deformation of the features. It has been shown by computer simulation that the new neocognitron recognizes characters more robustly than the conventional neocognitron.

However, several problems remain to be solved in the future. We used the same set of training patterns to train the new network as the one used to train the conventional neocognitron. We can expect a better performance if the training patterns are modified to fit the new network architecture.

We feel that the scale of the network tested in this paper is not large enough to bring the ability of the dual C-cell layer into full play. A much better performance could be expected if the density of the cells in the input layer is increased.

Although we concluded that dualization of C-cell layers of higher stages does not produce any significant effects, it might be still effective if an eccentric blurring is introduced. In the eccentric blurring, the position of the smallest blurring is not necessarily placed at the center of the receptive field of an S-cell, but is placed at the position of one of the local features constituting the global feature to be extracted by the S-cell. Creation of such eccentric blurring is another future problem to be tested.

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

Fukushima, K. (1989). Analysis of the process of visual pattern recognition by the neocognitron *Neural Networks*, **2**(6), 413–420.

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

Fukushima, K., & Wake, N. (1991). Handwritten alphanumeric character recognition by the neocognitron *IEEE Transactions on Neural Networks*, **2**(3), 355–365.

Hiroshige, K., Okada, M., & Fukushima, K. (1992) Neocognitron with a dual C-cell layer (in Japanese) *Technical Report, IEICE*, No. **NC91**-160.

LeCun, Y., Boser, B., Denker, J. S, et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, **1**, 541–551.