IJCAI-79

第6回人工知能国際会議

Proceedings of the Sixth International
Joint Conference on Artificial Intelligence
Tokyo, August 20–23, 1979
Volume One

IJCAI-79
Sponsored by: International Joint Conferences on Artificial Intelligence, Inc.
Information Processing Society of Japan

THE PROCEEDINGS WERE PRINTED WITH THE PARTIAL ASSISTANCE OF A GRANT FROM "THE COMMEMORATIVE ASSOCIATION FOR THE JAPAN WORLD EXPOSITION"

Cover design by J. Rindfleisch

SELF-ORGANIZATION OF A NEURAL NETWORK WHICH GIVES POSITION-INVARIANT RESPONSE

Kunihiko Fukushima

NHK Broadcasting Science Research Laboratories
1-10-11, Kinuta, Setagaya, Tokyo 157, Japan

In this paper, I propose a new algorithm for self-organizing a multilayered neural network which has an ability to recognize patterns based on the geometrical similarity of their shapes. This network, whose nickname is "neo-cognitron", has a structure similar to the hierarchy model of the visual nervous system proposed by Hubel and Wiesel. The network consists of a photoreceptor layer followed by a cascade connection of a number of modular structures, each of which is composed of two layers of cells connected in a cascade. The first layer of each module consists of "S-cells", which show characteristics similar to simple cells or lower order hypercomplex cells, and the second layer consists of "C-cells" similar to complex cells or higher order hypercomplex cells. The imput synapses to each S-cell have plasticity and are modifiable. The network has an ability of unsupervised learning: We don't need any "teacher" during the process of self-organization, and it is only needed to present a set of stimulus patterns repeatedly to the input layer. The network has been simulated on a digital computer. After completion of self-organization, the stimulus patterns has become to elicit their own response from the last C-cell layer. That is, the response of the last C-cell layer changes without fail, if a stimulus patterns of a different category is presented to the input layer. The response of that layer, however, is not affected by the pattern's position at all. Neither is it affected by a certain amount of changes of the pattern's shape or size.

1. INTRODUCTION

In this paper, I propose a new algorithm for self-organizing a neural network which has an ability to recognize stimulus patterns based on the geometrical similarity of their shapes regardless of their positions and slight distortions of their shapes.

This network is given a nickname "neocognitron", because it is a further extension of the "cognitron", which is a self-organizing multilayered neural network proposed by the author before [1]. Incidentally, the conventional cognitron also had an ability to recognize patterns, but its response was dependent upon the position of the stimulus pattern. That is, the same patterns which were presented at different positions were taken as different patterns by the conventional cognitron.

The neocognitron has also a multilayered structure. It has an ability of unsupervised learning: We don't need any "teacher" during the process of self-organization, and it is only needed to present a set of stimulus patterns repeatedly to the input layer. After completion of self-organization, the network acquires a structure similar to the hierarchy model of the visual nervous system proposed by Hubel and Wiesel [2], [3]. The response of the cells of the last layer of the network is dependent only upon the shape of the stimulus pattern, and is not affected by the position of pattern presentation. That is, the network has an ability to recognize patterns without affected by the position of the patterns.

2. STURUCTURE OF THE NETWORK

The network which is proposed here consists of a photoreceptor layer followed by a cascade-connection of a number of madular structures, each of which is composed of two layers of cells connected in a cascade. The first layer of each module consists of "S-cells", which correspond to simple cells or lower order hypercomplex cells according to the classification of Hubel and Wiwsel. The second layer of the module consists of "C-cells", which correspond to complex cells or higher order hypercomplex cells.

The cells in each layer are divided into many subgroups according to the optimum stimulus features of their receptive fields. Since the cells in each subgroup are set in a two dimensional array, we call the subgroup as a "cell-plane". It is assumed that all the cells in a single cell-plane have the same spatial distribution of the input synaptic connections, and only the position of the presynaptic cells are shifted in parallel from cell to cell. Hence, all the cells in a single cell-plane have receptive fields of the same functions, but at different positions.

Let $u_{S\ell}(k_\ell, n)$ be the output of an S-cell in the k_ℓ-th cell-plane in the ℓ-th module, and let $u_{C\ell}(k_\ell, n)$ be the output of a C-cell in that module, where n is two-dimensional co-ordinates representing the position of these cell's receptive fields in the input layer U_0. Besides these excitatory cells, we have inhibitory cells $v_{S\ell}(n)$ and $v_{C\ell}(n)$ in the module.

The output of these cells are represented as follows:

$$u_{S\ell}(k_\ell, n) = r_\ell \cdot \varphi\left[\frac{1+\sum_{k_{\ell-1}=1}^{K_{\ell-1}} \sum_{\nu \in S_\ell} a_\ell(k_{\ell-1}, \nu, k_\ell) \cdot u_{C\ell-1}(k_{\ell-1}, n+\nu)}{1+\frac{2r_\ell}{r_\ell}\cdot b_\ell(k_\ell)\cdot v_{C\ell-1}(n)} - 1\right] \tag{1}$$

$$\text{where} \quad \varphi[x] = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{2}$$

$$v_{C\ell-1}(n) = \sqrt{\sum_{k_{\ell-1}=1}^{K_{\ell-1}} \sum_{\nu \in S_\ell} c_{\ell-1}(\nu)\cdot u_{C\ell-1}^2(k_{\ell-1}, n+\nu)} \tag{3}$$

$$u_{C\ell}(k_\ell, n) = \psi\left[\frac{1+\sum_{\nu \in D_\ell} d_\ell(\nu)\cdot u_{S\ell}(k_\ell, n+\nu)}{1+v_{S\ell}(n)} - 1\right] \tag{4}$$

$$\text{where} \quad \psi[x] = \varphi\left[\frac{x}{1+x}\right] \tag{5}$$

$$v_{S\ell}(n) = \frac{1}{K_\ell}\sum_{k_\ell=1}^{K_\ell}\sum_{\nu \in D_\ell} d_\ell(\nu)\cdot u_{S\ell}(k_\ell, n+\nu) \tag{6}$$

Where, $a_\ell(k_{\ell-1}, \nu, k_\ell)$ and $b_\ell(k_\ell)$ represent the strength (efficiency) of the modifiable excitatory and inhibitory synapses, respectively. Parameter $c_{\ell-1}(\nu)$ and $d_\ell(\nu)$ represent the strength of the unmodifiable excitatory synapses. Positive parameter r_ℓ controls the intensity of the inhibitory input, and the large value of r_ℓ gives higher selectivity of the cell's response.

Let $u_0(n)$ be the output of a cell of the input layer U_0. In the above equations, it is assumed that $u_{C\ell-1}(k_{\ell-1}, n)$ stands for $u_0(n)$ in case of ℓ=1. Incidentally, $K_{\ell-1}$=1 for ℓ=1.

The inhibitory cell $v_{C\ell-1}(n)$ has an r.m.s.-type input-to-output characteristic as indicated by eq. (3). The employment of r.m.s.-type cells is useful for endowing the network with an ability to make reasonable evaluation of the similarity between the stimulus patterns. Its usefulness was analytically proved for a cognitron [4], and the same discussion can also be applied to our new network.

3. THE SELF-ORGANIZATION OF THE NETWORK

One of the fundamental hypotheses employed in this network is the assumption that all the S-cells in a single cell-plane have afferent synaptic connections of identical spatial distribution, and that only the position of the presynaptic cells shift in parallel in accordance with the S-cell's position. Therefore, in eq. (1), the modifiable synaptic connection $a_\ell(k_{\ell-1}, \nu, k_\ell)$ is assumed to be independent of the position n of the presynaptic cell $u_{S\ell}(k_\ell, n)$.

During the process of self-organization, several "representative" S-cells are selected every time when a stimulus pattern is presented. Here, we assume a certain mechanism which prohibit the selection of more than one representative from a single S-cell-plane. The detailed procedure for selecting the representatives are given later on.

Suppose a representative is selected from an S-cell-plane. The afferent synapses to the representative S-cells are reinforced with the same manner as in the case of r.m.s.-type cognitron [4], and the reinforcement of the synapses of all the other S-cells in that cell-plane follows those of their representative. These relations can be quantitatively expressed as follows.

Let cell $u_{S\ell}(\hat{k}_\ell, \hat{n})$ be selected as a representative. The modifiable synapses $a_\ell(k_{\ell-1}, \nu, \hat{k}_\ell)$ and $b_\ell(\hat{k}_\ell)$, which converge to the cells of the $\hat{k}_\ell$-th cell-plane, are reinforced by the amount shown below:

$$\Delta a_\ell(k_{\ell-1}, \nu, \hat{k}_\ell) = q_\ell \cdot c_{\ell-1}(\nu)\cdot u_{C\ell-1}(k_{\ell-1}, \hat{n}+\nu) \tag{7}$$

$$\Delta b_\ell(\hat{k}_\ell) = (q_\ell/2)\cdot v_{C\ell-1}(\hat{n}) \tag{8}$$

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

The cells in the S-cell-plane from which no representative is selected, however, do not have their input synapses reinforced at all.

The representatives are chosen by the following procedures. At first, we watch a group of

S-cells in a layer whose receptive fields are situated within a certain small area in the input layer. Each group contains cells from all the S-cell-planes. We call the group as an "S-column". There are a lot of such S-columns in a single layer. From each S-column, the cell which has yielded the largest output is chosen as a candidate for the representatives. In case only one candidate appears from an S-cell-plane, the candidate is unconditionally selected as the representative from that S-cell-plane. If more than two candidates appear from an S-cell-plane, however, only the one which has yielded the largest output among them is selected as the representative from that S-cell-plane. If no candidate appears from an S-cell-plane, no representative is selected from that S-cell-plane.

As it is seen from these discussions, if we consider that a single excitatory cell in the conventional cognitron [1] coresponds to a single S-cell-plane in the neocognitron, the procedures of reinforcement in the both systems are analogous to each other.

4. COMPUTER SIMULATION

The neural network proposed here has been simulated on a digital computer. In the computer simulation, we consider a seven layered network: $U_0 \to U_{S1} \to U_{C1} \to U_{S2} \to U_{C2} \to U_{S3} \to U_{C3}$. That is, the network has three stages of modular structures preceded by an input layer. The number of cell-planes in each layer is 24 for all the layers except U_0. The numbers of cells in these seven layers are 16×16, 16×16×24, 10×10×24, 8×8×24, 6×6×24, 2×2×24, and 24 from the front. In the last layer U_{C3}, each cell-plane contains only one cell, and that cell's receptive field covers the whole area of the input layer U_0.

We have presented five stimulus patterns, which are shown in the leftmost column in Fig. 1, repeatedly to the input layer U_0 of the network. The positions of presentation of these stimulus patterns have been randomely shifted at every presentation.

After a certain number of pattern presentations, each stimulus pattern has become to elicit an output only from one U_{C3}-cell, and conversely, this U_{C3}-cell has become selectively responsive only to that stimulus pattern. That is, none of the U_{C3}-cells responds to more than one stimulus pattern. It has been confirmed that the response of layer U_{C3} is not affected by the position of a stimulus pattern at all. Neither is it affected by the slight change of the shape or the size of the stimulus pattern.

Fig. 1 shows some examples of the stimulus patterns which the neocognitron has correctly recognized: All the stimulus patterns in each row of Fig. 1 have elicited the same response to U_{C3}-cells. That is, the response of layer U_{C3} is affected neither by shift of position like (a)-(c), nor by distortion of shape or size like (d)-(f), nor by some insufficiency of the pattern or some noise like (g).

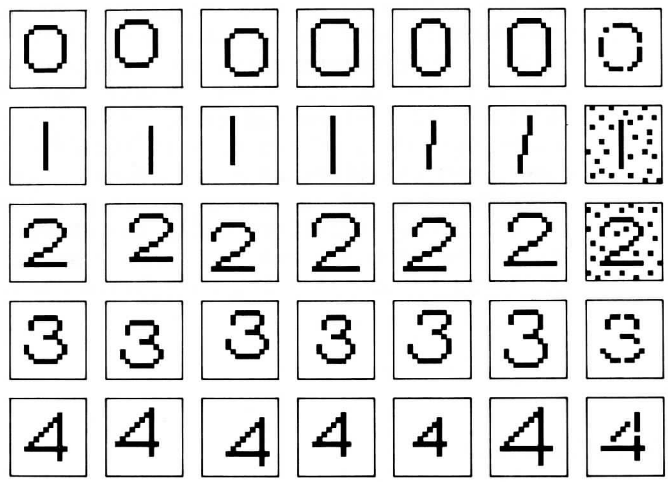

Fig. 1 Some examples of distorted stimulus patterns which the neocognitron has correctly recognized.

REFERENCES

[1] Fukushima, K. "Cognitron: A Self-organizing Multilayered Neural Network". Biol. Cybernetics 20:3/4 (1975) 121-136.

[2] Hubel, D.H., Wiesel, T.N. "Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex". J. Physiol. (Lond.) 160:1 (1962) 106-154.

[3] Hubel, D.H., Wiesel, T.N. "Receptive Fields and Functional Architecture in Two Nonstriate Visual Area (18 and 19) of the Cat". J. Neurophysiol. 28 (1965) 229-289.

[4] Fukushima, K. "Improvement in Pattern-Selectivity of a Cognitron" (in Japanese). Paper of Tech. Group, MBE78-27, Inst. Electronics Commun. Engrs. Japan, July 1978.