Pattern Recognition System with Top-Down Process of Mental Rotation

Shunji Satoh[1], Hirotomo Aso[1], Shogo Miyake[2], and Jousuke Kuroiwa[3]

[1] Department of Electrical Communications, Tohoku University
Aoba-yama05, Sendai 980-8579, JAPAN
{shun,aso}@aso.ecei.tohoku.ac.jp
http://www.aso.ecei.tohoku.ac.jp/~shun/index-e.html
[2] Department of Applied Physics, Tohoku University
Aoba-yama04, Sendai 980-8579, Japan
miyake@nlap.apph.tohoku.ac.jp
[3] The Division of Mathematical and Information Sciences,
Hiroshima University, Higashi-Hiroshima 739-8521, Japan
jou@mis.hiroshima-u.ac.jp

Abstract. A new model which can recognize rotated, distorted, scaled, shifted and noised patterns is proposed. The model is constructed based on psychological experiments in a mental rotation. The model has two types of processes: (i) one is a bottom-up process in which pattern recognition is realized by means of a rotation-invariant neocognitron and a standard neocognitron and (ii) the other is a top-down process in which a mental rotation is executed by means of a model of associative recall in visual pattern recognition. In computer simulations, it is shown that the model can recognize rotated patterns without training those patterns.
Keywords. rotation-invariant neocognitron, rotated pattern, mental rotation, top-down process, pattern recognition

1 Introduction

In usual we can recognize patterns even if those are affected by distortions in shapes, shifts in positions, and/or rotations. However it is difficult to realize a recognition system which is robust for such many kinds of deformation. Actually, some of systems [1] can recognize rotated patterns but can not recognize distorted patterns, a system proposed in [2] cause a combinational explosion of units, and a system proposed in [3] is robust only for translation and rotations at every 90°. However, they are not necessarily satisfactory in the sense that they can not recognize rotated, distorted, shifted, scaled and noised patterns. For example, some models can recognize distorted patterns but can not recognize rotated ones. It is very important to realize a reliable recognition system which is insensitive to most deformations described above.

We have proposed a rotation-invariant neocognitron [4, 5], which is an extended model of a (standard) neocognitron [6], to satisfy most all demands of pattern recognition. It has been shown that the rotation-invariant neocognitron is robust for rotation, distortion, shift, scaling, and noise by computer simulations. The rotation-invariant neocognitron is constructed as a feedforward network based on the idea that simple patterns could be recognized by means of a bottom-up process even if the patterns are rotated, shifted, scaled and so on. The rotation-invariant neocognitron, however, learns implicitly rotated patterns in any angles when a standard (non-rotated) pattern is presented at a training phase. This learning process of the rotation-invariant neocognitron may be effective for simple patterns but not be effective for complicated patterns, *e.g.* Chinese characters, because huge number of cells/connections and very long learning time are needed to make possible to recognize such rotated and complicated patterns. Although humans can correctly recognize rotated patterns, it does not necessarily mean that humans learn or store all rotated patterns in any angles.

We think that such complicated and rotated patterns could not be recognized by a bottom-up process only, and top-down processes could play one of important roles in order to realize a complicated and rotated pattern recognition. Indeed, psychological experiments have been shown that information of mental rotation play one of important roles in recognizing rotated patterns. Shepard and Metzler showed by psychological experiments that larger rotation angles of a pattern causes longer recognition time of the pattern [7]. In their study, they proposed the existence of *mental rotation* in recognition processes in such a rotated pattern. We can also see from the result that humans execute a mental rotation after a tentative result of recognition of the rotated pattern is obtained (see Fig. 1(b)), because humans can not calculate the difference of directions between the rotated pattern and a standard one (Fig. 1(d)) without the tentative result for the rotated one. Here, the "tentative result" means that the result is not necessarily coincide with a correct category of a presented pattern. To know the direction of the rotated pattern, humans have to make a hypothesis (tentative result) for the category of the rotated pattern. We consider that mental rotation has an information in verifying the hypothesis, and the recognition processes of rotated patterns are done through the verifications of iterative hypotheses by means of the information of a mental rotation and knowledges of stored patterns [8]. Therefore, we assume that a mental rotation is a top-down process. On the contrary, it is apparent that a recognition sequence by the neocognitron-type network is a bottom-up process.

In this paper, we examine the process of recognition of rotated patterns from a psychological point of view and construct a new model in which the mental rotation is executed by means of feedback processes (top-down processes). The aim of this paper, however, is not to construct a model which realize a mental rotation but to construct a more realistic and reliable model to recognize rotated, scaled, distorted shifted and noised patterns.

2 Algorithm of the New Model

From the study for recognition of complicated and rotated patterns discussed in the previous section, we give an algorithm to recognize complicated and rotated patterns in Fig. 1. Next, we will give more details about processes or structures shown in Fig. 1.

2.1 Bottom-up Process: Hybrid-Neocognitron

Basic Idea to Construct a Bottom-up Process. We give a basic idea to construct a bottom-up recognition system, which is able to recognize distorted, scaled, shifted and noised patterns.

If we recognize a rotated pattern, what kinds of patterns do we use an information of mental rotation? We classify patterns into two classes, *mental-rotation-required* and *mental-rotation-free* patterns, by use of the criterion whether the mental rotation is needed to *correctly* recognize a rotated pattern or not. Humans can correctly recognize rotated patterns

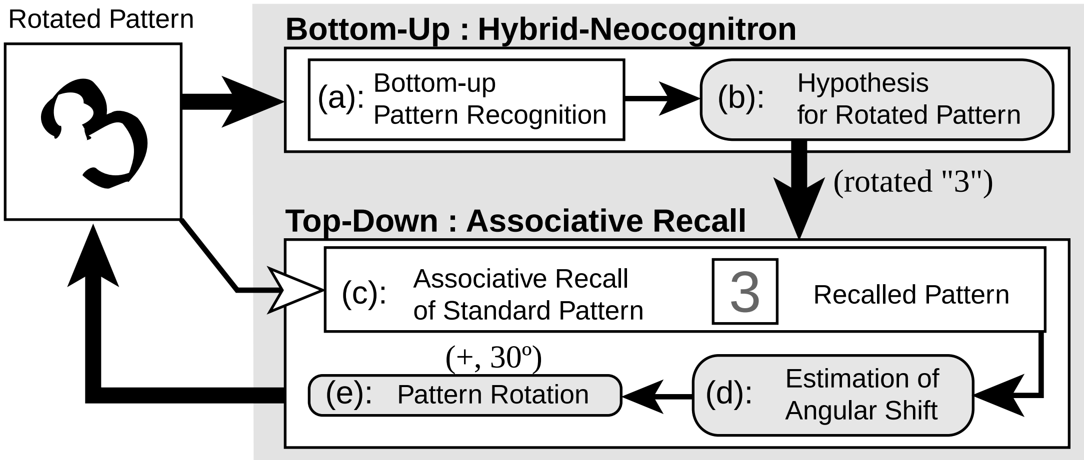
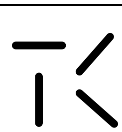
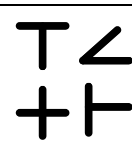
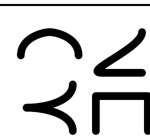


Fig. 1. Overview of the system for recognition of complicated and rotated patterns. The process is divided into two parts: (i) a bottom-up process of a recognition system and (ii) a top-down process of a system with associative recall. Mental rotation, (d) and (e), requires a tentative recognition result (b) and the recalled pattern (c).

and those orientations in an instant without help of mental rotation if those rotated patterns are comparatively simple like a T-shaped pattern, cross-shaped pattern or ∠-shaped pattern. Such simple patterns are composed of two or more line segments with different orientations (we call such segments oriented segments [10]). We classify those patterns into mental-rotation-free patterns. On the other hand, rotated numerals, alphabets or Chinese characters are *not correctly* recognized in an instant and those orientations are not determined in an instant. Those complicated patterns are composed of many T-shaped patterns, cross-shaped patterns and so on. We classify such patterns into mental-rotation-required patterns. Examples of each classes in numerals are depicted in Table. 1. The difference be-

Table 1. The classification of complexities of patterns from a psychological point of view and complexities to be detected in each module of a rotation-invariant neocognitron or a standard neocognitron. The lower layer detects comparatively simple patterns (mental-rotation-free patterns) and the higher layer detects complicated patterns (mental-rotation-required patterns).

	module			
	U_1	U_2	U_3	U_4
Features				2 3 5 7
Class	mental-rotation-free		mental-rotation-required	

tween mental-rotation-free and mental-rotation-required patterns is due to whether humans have already learned and stored all rotated patterns in any angles or not.

In this paper, we construct the bottom-up system for recognition of complicated and rotated patterns based on the system of a rotation-invariant neocognitron and a standard neocognitron [6]. In the rotation-invariant neocognitron and the standard neocognitron complexities of patterns can be classified into four classes. The rotation-invariant neocognitron and the standard neocognitron are multi-layered networks and each of those has the retina, U_0, and basically four modules U_1, U_2, U_3 and U_4 (each model has nine layers since each module $U_l, l = 1, 2, 3, 4$, consists of two layers, $U_{\mathrm{S}l}$ and $U_{\mathrm{C}l}$). Lower modules, U_1 and U_2, detect comparatively simple patterns and higher modules, U_3 and U_4, complicated patterns. Examples of patterns to be detected in each module is depicted in Table 1. We can regard patterns detected in U_1 and U_2 as mental-rotation-free patterns because those patterns are oriented segments or composed of two or more oriented segments and patterns detected in U_3 and U_4 as mental-rotation-required patterns because those require a mental rotation for *correct* recognition.

We can see from Table 1 that the psychological classification of complexities of patterns corresponds to the classification by the serial number of modules in the rotation-invariant neocognitron or the standard neocognitron. Therefore, the model of a bottom-up type recognition system have two modules: (i) the lower modules are composed of ones of a rotation-invariant neocognitron, in which all the simple rotated patterns in any angles (mental-rotation-free patterns) are detected, and (ii) the higher modules are composed of a standard neocognitron, in which only standard complicated patterns are detected.

The schematic structure is depicted in the upper part of Fig. 2. The number of cells is reduced by replacing higher modules of a rotation-invariant neocognitron by higher modules of a standard neocognitron. We name the bottom-up system a hybrid-neocognitron. We note that the hybrid-neocognitron is able to recognize distorted, scaled, shifted and noised patterns since the hybrid-neocognitron inherits the all functions of a standard neocognitron.

Tentative Recognition of Rotated Patterns Using Hybrid-Neocognitron If those values of threshold of feature-detecting cells and other parameters remain fixed in the hybrid-neocognitron, the model can not recognize rotated patterns (all cells in $U_{\mathrm{C}4}$, gnostic cells, make no response) because a standard neocognitron can not recognize rotated patterns. However, in order to produce an information of a mental rotation in the rotated pattern (see Fig. 1(b)), the hybrid-neocognitron, however, has to give a tentative recognition result in some way. We note again that the tentative result of recognition needs not to be correct one, and the tentative result is a hypothesis produced by a hybrid-neocognitron. Moreover, a mental rotation gives an information in verifying the hypothesis. The problem is that how to make a tentative recognition result for the rotated pattern using a bottom-up model only which has not learned or stored rotated patterns.

Here, we give results of a psychological experiment with respect to recognition of rotated patterns by twelve subjects. The aim of the experiment is to analyze whether subjects can make a correct recognition in rotated patterns *without* a mental rotation or not. The conditions is

- an experimenter notes a subject that a pattern is rotated,
- a pattern shown in Fig. 3 is presented in the front of the subject,
- and the subject is forced to report the category of the pattern *as soon as possible*.

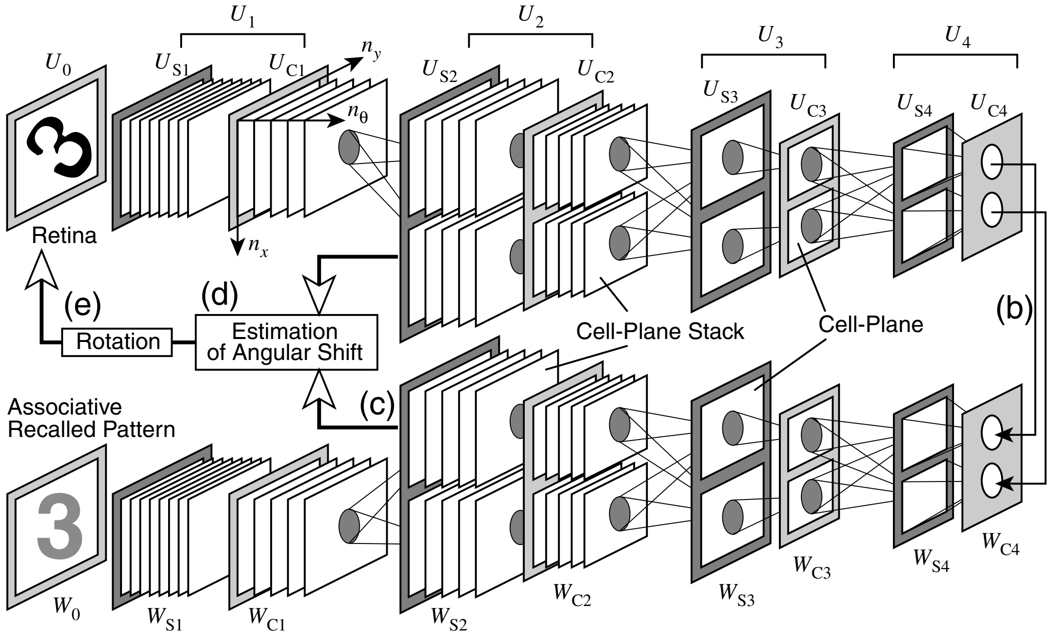


Fig. 2. Upper part: the structure of a bottom-up system, a hybrid-neocognitron. Lower part: the structure of a model of associative recall. The hybrid-neocognitron is composed of a rotation-invariant neocognitron in the lower module and a standard neocognitron in the higher module. The structure of associative recall is same as one of the hybrid-neocognitron. (b)–(e) are correspond to processes (b)–(e) shown in Fig. 1.

The last item of the condition intends to force a subject to recognize a rotated pattern without a mental rotation. The result is very interesting. All of twelve subjects report that "the pattern is 2" beyond doubt when the leftmost pattern in Fig. 3 is presented although the pattern is a rotated one of a mirror image of "2". We conclude that the absolute positions of local features which compose the whole of a pattern is ignored if one has to recognize the pattern using only a bottom-up-type process. In our bottom-up system, these results are given by extending the region of blurring operation by C-cells in U_{C2}, U_{C3} and U_{C4} and by decreasing threshold values of S-cells in U_{S3} and U_{S4}. Thus, our bottom-up system with the extending blurring operation and decreasing threshold values can give a tentative recognition in a rotated pattern. Details are explained in the Section 3.

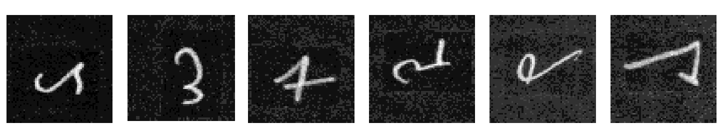


Fig. 3. Examples of patterns to test a recognition ability without mental rotation.

2.2 Associative Recall of Standard Pattern

After making hypothesis for a rotated pattern, the standard pattern is recalled using the knowledge about the pattern. The associative recall is given by the model which is constructed based on the selective attention and associative recall model given by Fukushima [11]. In the present paper the function of selective attention is omitted for simplicity.

The structure and connections of our associative recall model are similar to the one of the hybrid-neocognitron but the flow of information is opposite directions, from the highest module to the lowest module (see the lower part of Fig. 2). As the result, a recalled pattern is a blurred pattern (an average pattern) of the tentative recognition result for a rotated pattern.

3 Mathematical description of the Model

3.1 Bottom-up Process: Hybrid-Neocognitron

Mathematical description of the hybrid-neocognitron is same as the one of the rotation-invariant neocognitron because the rotation-invariant neocognitron includes a standard neocognitron.

The structure of the hybrid-neocognitron is shown in the upper part of Fig. 2. U_0 is a model of retina. $U_{\mathrm{S}l}$ denotes a layer consisting of S-cells in the lth module, and $U_{\mathrm{C}l}$ a layer consisting of C-cells. An S-cell is a feature-detecting cell and a C-cell has a blurring function so that the model is robust for rotation or distortion. Lower layers, $U_{\mathrm{S}1}$, $U_{\mathrm{C}1}$, $U_{\mathrm{S}2}$ and $U_{\mathrm{C}2}$, is composed of a number of cell-plane stacks [4, 5], and different cell-plane stacks detect different features of inputs (a cell-plane stack is referred to as “CPS”). A CPS is composed of a number of cell-planes (referred to as CP(s)) [6], and each CP detects a different rotation angle of the local pattern (a local pattern which composes an input pattern is called as “features”). Each cell in a CPS is located in a three-dimensional space. A rotational information of a local pattern is represented by the serial number n_θ assigned to a CP in a CPS, and the positional information by the position (n_x, n_y) of a firing cell in the specific CP.

An output response of an S-cell located on $\boldsymbol{n} = (n_x, n_y, n_\theta)$ of the kth CPS in the lth module is denoted by $u_{\mathrm{S}l}(\boldsymbol{n}, k)$, and an output response of a C-cell by $u_{\mathrm{C}l}(\boldsymbol{n}, k)$. The output response $u_{\mathrm{S}l}(\boldsymbol{n}, k)$ is given by

$$u_{\mathrm{S}l}(\boldsymbol{n}, k) = r_l \cdot \phi\left[\frac{1+e}{1+\dfrac{r_l}{1+r_l}\cdot i} - 1\right], \tag{1}$$

$$\phi\left(x\right) = \max\left(x, 0\right), \tag{2}$$

where

$$e = \sum_{\kappa=0}^{K_{\mathrm{C}l-1}-1} \sum_{\boldsymbol{\nu}\in A_l} a_l(\boldsymbol{\nu}, \boldsymbol{n}, \kappa, k) \cdot u_{\mathrm{C}l-1}(\boldsymbol{n} \underset{T_{\mathrm{C}l-1}}{\oplus} \boldsymbol{\nu}, \kappa), \tag{3}$$

$$i = b_l(k) \cdot u_{\mathrm{V}l}(\boldsymbol{n}). \tag{4}$$

A binomial operator $\underset{M}{\oplus}$ with M is defined by

$$\begin{cases} n_x \underset{M}{\oplus} \nu_x \overset{\text{def}}{=} n_x + \nu_x, \\ n_y \underset{M}{\oplus} \nu_y \overset{\text{def}}{=} n_y + \nu_y, \\ n_\theta \underset{M}{\oplus} \nu_\theta \overset{\text{def}}{=} (n_\theta + \nu_\theta) \bmod M. \end{cases} \tag{5}$$

Here r_l denotes a threshold value of an S-cell, $a_l(\boldsymbol{\nu}, \boldsymbol{n}, \kappa, k)$ represents an excitatory connection from a C-cell to an S-cell and $b_l(k)$ an inhibitory connection from a V-cell to an S-cell —a V-cell sends inhibitory inputs to S-cells and are not depicted in Fig. 2 for simplicity. Each connection is linked to a restricted number of C-cells in the preceding module, $A_l \subset \boldsymbol{Z}^3$, $K_{\mathrm{C}l}$ denotes the number of CPSs in the $U_{\mathrm{C}l}$ layer, and $T_{\mathrm{C}l}$ the number of CPs in the $U_{\mathrm{C}l}$ layer.

An output response of a V-cell is given by

$$u_{\mathrm{V}l}(\boldsymbol{n}) = \sqrt{\sum_{\kappa=0}^{K_{\mathrm{C}l-1}-1} \sum_{\boldsymbol{\nu} \in A_l} c_l(\boldsymbol{\nu}) \cdot \left\{ u_{\mathrm{C}l-1}(\boldsymbol{n} \underset{T_{\mathrm{C}l-1}}{\oplus} \boldsymbol{\nu}, \kappa) \right\}^2} \tag{6}$$

where $c_l(\boldsymbol{\nu})$ is an excitatory connection from a C-cell to a V-cell, which takes a fixed value during a learning.

An output response of a C-cell is given by

$$u_{\mathrm{C}l}(\boldsymbol{n}, k) = \psi \left[\sum_{\boldsymbol{\nu} \in D_l} d_l(\boldsymbol{\nu}) \cdot u_{\mathrm{S}l}(\boldsymbol{n} \underset{T_{\mathrm{S}l}}{\oplus} \boldsymbol{\nu}, k) \right], \tag{7}$$

where the function ψ is defined by

$$\psi(x) = \frac{\phi(x)}{\phi(x) + 1}. \tag{8}$$

Here $d_l(\boldsymbol{\nu})$ is an excitatory connection from an S-cell to a C-cell, $D_l \subset \boldsymbol{Z}^3$ represents a restricted region of the connection and $T_{\mathrm{S}l}$ the number of CPs in the $U_{\mathrm{S}l}$ layer. C-cells in the last layer, $u_{\mathrm{C}4}(\boldsymbol{n}, k)$, are gnostic cells, and so a maximumly firing C-cell represents a recognition result with highest likelihood among the all of categories.

The parameters about the number of CPSs, CPs and cells are given in Table 2. The initial sizes of the areas, A_l and D_l, are given in Table 3.

Table 2. The number of cell-plane stack, and the number of cells in one cell-plane stack. The numbers in columns marked by asterisks are not defined, and the numbers in parentheses in columns are ones after completion of learning.

	$l=0$	$l=1$	$l=2$	$l=3$	$l=4$		$l=0$	$l=1$	$l=2$	$l=3$	$l=4$
$K_{\mathrm{S}l}$	$*$	1	(6)	(22)	(10)	$K_{\mathrm{C}l}$	1	1	(6)	(22)	(10)
$N_{\mathrm{S}l}$	$*$	59	17	13	3	$N_{\mathrm{C}l}$	61	19	17	10	1
$T_{\mathrm{S}l}$	$*$	16	8	1	1	$T_{\mathrm{C}l}$	1	8	4	1	1

During a learning phase, excitatory connections $a_l(\boldsymbol{\nu}, \boldsymbol{n}, \kappa, k)$ and inhibitory connections $b_l(k)$ are modified according to an unsupervised learning using seed-cell [9]. An algorithm proposed by authors, auto-generating algorithm [10], by which learning time is drastically reduced in the unsupervised learning is also adopted.

During a recognition phase, all gnostic cell would make no response if a largely rotated pattern is presented in U_0. At that time, the hybrid-neocognitron execute the following two processes; (i): the model decreases the value of threshold of S-cells in $U_{\mathrm{S}3}$ and $U_{\mathrm{S}4}$ and (ii):

Table 3. The size of area connected with one cell, A_l and D_l. The size in parentheses in columns is one after completion of extending the size, discussed in section 2.1.

	$l=1$	$l=2$	$l=3$	$l=4$
A_l	$3\times 3\times 1$	$3\times 3\times 8$	$5\times 5\times 1$	$3\times 3\times 1$
D_l	$5\times 5\times 3$	$3\times 3\times 3(5\times 5\times 8)$	$5\times 5\times 1(7\times 7\times 1)$	$3\times 3\times 1(5\times 5\times 1)$

spreads the blurring region of C-cells in U_{C2}, U_{C3} and U_{C4} as discussed in the section 2.1. The process (i) is denoted by $r_l := \gamma\cdot r_l (l = 3,4)$, where the notation := is used in the sense of computer language PASCAL and γ is an attenuation constant. Extended regions of blurring operation is given in parentheses of Table 3.

3.2 Top-Down Process: Associative Recall

The structure of the model of associative recall (the lower part in Fig. 2) is same as the one of hybrid-neocognitron except the direction of the flow of information. Each layer is denoted by W_0, W_{Sl} and W_{Cl}, $(l = 1,2,3,4)$. The output of the recognition layer in a hybrid-neocognitron, U_{C4}, is sent back to lower modules through backward paths of the associative recall model and the flow reaches the recall layer W_0.

The output of a w_{Cl} cell in W_{Cl} and the cell w_{Vl} in W_{Sl} in the backward paths are given by

$$w_{Cl}(\boldsymbol{n},k) = \psi\left[\alpha_l\cdot\left\{\sum_{\kappa=1}^{K_{Sl+1}}\sum_{\boldsymbol{\nu}\in A_{l+1}} a_{l+1}(\boldsymbol{\nu},\boldsymbol{n},k,\kappa)\cdot w_{Sl+1}(\boldsymbol{n}\underset{T_{Sl+1}}{\ominus}\boldsymbol{\nu},\kappa) - \sum_{\boldsymbol{\nu}\in A_{l+1}} c_{l+1}(\boldsymbol{\nu})\cdot w_{Vl+1}(\boldsymbol{n}\underset{T_{Sl+1}}{\ominus}\boldsymbol{\nu})\right\}\right], \tag{9}$$

$$w_{Vl+1}(\boldsymbol{n}) = \frac{r_{l+1}}{1+r_{l+1}}\sum_{\kappa=1}^{K_{Sl+1}} b_{l+1}(\kappa)\cdot u_{Sl+1}(\boldsymbol{n},\kappa), \tag{10}$$

where α_l is a positive constant. The output of a w_{Sl} cell in W_{Sl} is given by

$$w_{Sl}(\boldsymbol{n},k) = \beta_l\cdot\sum_{\boldsymbol{\nu}\in D_l} d_l(\boldsymbol{\nu})\cdot w_{Cl}(\boldsymbol{n}\underset{T_{Cl}}{\ominus}\boldsymbol{\nu},k), \tag{11}$$

where β_l is a positive constant.

The learning of the model of associative recall is executed by following the learning of the hybrid-neocognitron. For example, if a new CPS or CP is generated in the hybrid-neocognitron, the same CPS or CP is also generated in the associative recall model. The network is designed so that the variable backward connections is automatically reinforced in the following manner: after finishing the reinforcement of the forward connections, the backward connections descending from a w_S cell are automatically reinforced to have a strength proportional to the forward connections ascending to the u_S cell which makes a pair with the w_S cell.

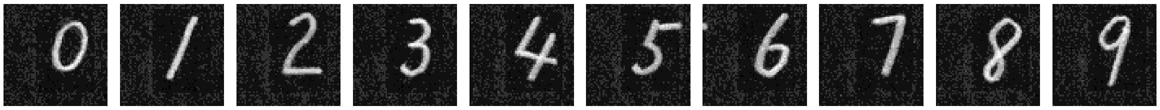


Fig. 4. Examples of training patterns.

3.3 Angle Estimation

The difference in orientation between a rotated pattern and the recalled standard pattern is estimated in order to determine the angle of a mental rotation. In estimating the difference, as shown by the psychological experiment in Section 2.1 the absolute positional information is ignored at comparing those two patterns. The angle estimator compares projected patterns on n_θ-axes in $U_{\mathrm{S}l}$ and $W_{\mathrm{S}l}$ because the absolute positional information of a feature is represented by a position (n_x, n_y). The difference of orientations, Θ, is given by

$$\Theta = \mathop{\mathrm{argmax}}_{\Delta\theta=0,\pm1,\cdots,\pm T_{\mathrm{S2}}-1} \left(u_{\mathrm{S2}}\left(n_\theta, k\right) \cdot w_{\mathrm{S2}}(n_\theta \underset{T_{\mathrm{S2}}}{\oplus} \Delta\theta, k) \right), \tag{12}$$

$$u_{\mathrm{S2}}\left(n_\theta, k\right) = \sum_{n_x, n_y \in N_{\mathrm{S2}}} u_{\mathrm{S2}}\left(\boldsymbol{n}, k\right), \quad w_{\mathrm{S2}}\left(n_\theta, k\right) = \sum_{n_x, n_y \in N_{\mathrm{S2}}} w_{\mathrm{S2}}\left(\boldsymbol{n}, k\right). \tag{13}$$

4 Simulation

We examine the ability of the new model for the recognition of realistic hand-written numerical patterns provided by ETL-1 database [1]. We use five hundreds training patterns (see Fig. 4 for examples) and one thousands test patterns. The set of test patterns includes patterns which are generated by rotating standard patterns in ETL1 database by use of a computer. Examples of correctly recognized patterns is shown in Table 4. The value in column Θ^{org} in Table 4 is an actual rotational difference between a pattern in the database and the rotated ones. The difference of angles between columns is due to the resolution in orientation of the hybrid-neocognitron. The resolution is $45°$ since comparison of a standard pattern and a rotated one is executed in U_{S2} and W_{S2} and orientation is quantized into $T_{\mathrm{S2}} = 8$ direction in those layers. The error is not significant and does not influence the ability of recognition because hand-written patterns change the shape by writers and one can not estimate precise rotational angles of those patterns. Recognition rate for non-rotated test patterns is 91.0% and rotated test patterns is 90.5% in spite of the fact that the model does not learn or not store rotated patterns. This result shows the effectiveness of the model.

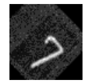
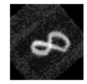
When only the hybrid-neocognitron is used to recognize test patterns, the recognition rate for rotated patterns is 16.6%. It turns out that the model of mental rotation is effective for recognition of rotated pattern in the case that a bottom-up system does not learn or not store rotated patterns.

5 Conclusion

We construct a new model for pattern recognition which is robust for rotation as well as distortion, scaling, shift in position and noise. The new model is designed so that the process

[1] A character database published from the Electrotechnical Laboratory, Japan.

Table 4. Samples of correctly recognized patterns. Θ^{org} is rotational angle from the standard pattern. Θ is an angle estimated by the angle estimator.

test pattern										
Θ^{org}	$+90°$	$+90°$	$+70°$	$-80°$	$-50°$	$+90°$	$-30°$	$-45°$	$-50°$	$-30°$
Θ	$+90°$	$+90°$	$+90°$	$-90°$	$-45°$	$+90°$	$-45°$	$-45°$	$-45°$	$-45°$

includes the result of psychological experiments. By computer simulations the effectiveness of the new model is shown by using a numerical hand-written database. We intend to improve the recognition rate by adding an additional module used in a standard neocognitron, and expect that the model will show a high recognition rate for realistic patterns.

References

1. M. Fukumi, S. Omatu, and Y. Nishikawa, "Rotation-Invariant Neural Pattern Recognition System Estimating a Rotation Angle," *IEEE Trans., Neural Network*, **Vol. 8**, pp. 568–581, 1997.
2. M. B. Reid, L. Spirkovska, and E. Ochoa, "Rapid training of higher order neural networks for invariant pattern recognition," *Proc. Int. Joint Conf. Neural Networks*, **Vol. 1**, pp. 689–692, 1989.
3. B. Widrow, R. G. Winter and R. A. Baxter, "Layered neural nets for pattern recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, **Vol. 36**, pp. 1109–1118, 1988.
4. S. Satoh, J. Kuroiwa, H. Aso and S. Miyake, "Recognition of rotated patterns using neocognitron," *Proc. Int. Conf. Neural Information Processing*, **Vol. 1**, pp. 112–116, 1997.
   `http://www.aso.ecei.tohoku.ac.jp/~shun/postscript/iconip97.ps.gz`
5. S. Satoh, J. Kuroiwa, H. Aso and S. Miyake, "A rotation-invariant Neocognitron (in Japanese)," *IEICE Trans.*, **Vol.J81-DII**, 1998.
6. K. Fukushima, "Neocognitron : A hierarchical neural network capable of visual pattern recognition," *Neural Networks*, **Vol. 1**, No. 2, pp. 119–130, 1988.
7. R.N.Shepard and J. Metzler, "Mental rotation of three-dimensional object," *Science*, **Vol. 171**, pp. 701–703, 1971.
8. C. Stanfill and D. Waltz, "Toward memory-based reasoning," *Communication ACM*, **Vol. 29**, pp. 1213–1228, 1986.
9. K. Fukushima and N. Wake, "An improved learning algorithm for the neocognitron," *Proc. of the Int. Conf. on Artificial Neural Networks*, pp. 4–7 (1992).
10. S. Satoh, J. Kuroiwa, H. Aso and S. Miyake, "Recognition of hand-written patterns by rotation-invariant neocognitron," *Proc. Int. Conf. Neural Information Processing*, **Vol. 1**, pp. 295–299, 1998.
    `http://www.aso.ecei.tohoku.ac.jp/~shun/postscript/iconip98.ps.gz`
11. K. Fukushima, "Neural network model for selective attention in visual pattern recognition and associative recall," *Applied Optics*, **Vol.26**, pp. 4985–4992, 1987.