FIGURE ALPHABET HYPOTHESIS INSPIRED NEURAL NETWORK RECOGNITION MODEL

Ryoji Ohira†,‡, *Kenji Saiki*† *and Tomoharu Nagao*†

† Department of Information Media and Environment
Graduate School of Environment and Information Sciences
Yokohama National University
79-7, Tokiwadai, Hodogaya-ku, Yokohama, Kanagawa, 240-8501, Japan

‡ Mechanics Platform Systems Development Department
Systems Hardware Development Division Systems Hardware Company
Oki Electric Industry Co., Ltd.
3-1 Futaba-cho, Takasaki-shi, Gunma, 370-8585,Japan
E-mail: ohira@nlab.sogo1.ynu.ac.jp

ABSTRACT

The object recognition mechanism of human being is not well understood yet. On research of animal experiment using an ape, however, neurons that respond to simple shape（e.g. circle, triangle, square and so on） were found. And Hypothesis has been set up as human being may recognize object as combination of such simple shapes. That mechanism is called Figure Alphabet Hypothesis, and those simple shapes are called Figure Alphabet.

As one way to research object recognition algorithm, we focused attention to this Figure Alphabet Hypothesis. Getting idea from it, we proposed the feature extraction algorithm for object recognition. In this paper, we described recognition of binarized images of multifont alphabet characters by the recognition model which combined three-layered neural network in the feature extraction algorithm. First of all, we calculated the difference between the learning image data set and the template by the feature extraction algorithm. The computed finite difference is a feature quantity of the feature extraction algorithm. We had it input the feature quantity to the neural network model and learn by backpropagation (BP method).

We had the recognition model recognize the unknown image data set and found the correct answer rate. To estimate the performance of the contriving recognition model, we had the unknown image data set recognized by a conventional neural network.

As a result, the contriving recognition model showed a higher correct answer rate than a conventional neural network model. Therefore the validity of the contriving recognition model could be proved. We'll plan the research a recognition of natural image by the contriving recognition model in the future.

Keywords: figure alphabet hypothesis, neural network, feature extraction algorithm

1. INTRODUCTION

The clarification and modeling of the information processing mechanism about which the human being is excellent is extensively researched. Application to the image processing and the image recognition is expected. It has been clarifying how human does recognize the object from a physiologic research that used the monkey belonging to Macaca type (Japanese Macaque and Macaca Mulatta, etc.). However, because the processing of information on the brain has not been clarified, modeling is impossible.

In the such status, hypothesis of the recognition mechanism of the human being is reported[1][2]. The recognition of the human being is processed at the dorsal visual pathway and ventral visual pathway. The ventral visual pathway is the path that human recognizes the object, and the information is transmitted in this order: V1 (Primary Visual Cortex), V2, V4, TEO and TE area. In V1, the neuron that reacts to the line segment of a specific inclination exists and it is a column structure. It is clarified that the number of neurons, which react to the line segment of specific inclination, decrease as the route goes. On the other hand, from a physiologic research, it is clarified that the number of neurons which only react to specific shape and the complex figure at a middle level show up.

The characteristic of this neuron is called “Figure Feature Selectivity”[3][4]. In addition, the neuron which reacts to the resembling diagram is collected and becomes a column[5]. And Hypothesis has been set up as human being may recognize object as combination of such simple shapes. That mechanism is called “Figure Alphabet Hypothesis”,

and those simple shapes are called “Figure Alphabet”. However, shape and the number of Figure Alphabet have not been yet clarified[6].

The purpose of the this research is to make the recognition model which considered the pattern recognition mechanism of the human being. And, we are expecting the object of a natural image to be recognized. As one way to research object recognition algorithm, we focused attention to this Figure Alphabet Hypothesis. Getting idea from it, we contriving the feature extraction algorithm for object recognition[7]. The contriving feature extraction algorithm assumes the Figure Alphabet to be a simple dot pattern. And, the dot pattern is generated by using Genetic Algorithm (GA).

In this research, we contriving recognition model by whom three layer neural network is united to the feature extraction algorithm. And, we show the effectiveness of the recognition model designed by using the multifont image. This paper consists of five sections. Section 2 describes contriving recognition model. Section 3 describes experiment setting. Section 4 describes experiment result. Section 5 is describes the conclusions.

2. FIGURE ALPHABET RECOGNITION MODEL

2.1. Composition of the recognition model

The contriving recognition model is the recognition model which got an idea to the Figure Alphabet hypothesis. Figure 1 is a contriving recognition model. We supposed that the Figure Alphabet on the TE area is the feature extraction part of the simple dot pattern. And, we thought that the processing to recognize in the combination of the Figure Alphabet is the composition of the neural network.

The contriving recognition model is the structure which combined this feature extraction part and a neural network, and the dot pattern and the entry layer support. That is, when the number of the dot patterns is 10, the entry layer becomes 10.

2.2. Operation of the recognition model

The operation of the contriving recognition model is described. First, it is the way of generating a dot pattern in GA. The parameter of the feature extraction part is shown in Table 1. We input a learning image data set to GA. GA generates a dot pattern in the initial generation and generates a best dot pattern in repeating crossing, selection, a mutation.

Next, we have the neural network learning in the learning image. The parameter of the neural network is shown in Table 2. At first, we computed finite difference between the learning image and the dot pattern by the feature extraction algorithm. The computed finite difference is the feature quantity of the feature extraction algorithm and is the entry of the neural network. We input the feature quantity to the neural network and make learn it by the backpropagation (BP method).

After that, we make a recognition model recognize a unknown image. We input unknown image to the feature extraction part and make compute a finite difference with the dot pattern. The feature quantity is input to the neural network and a recognition result is output.

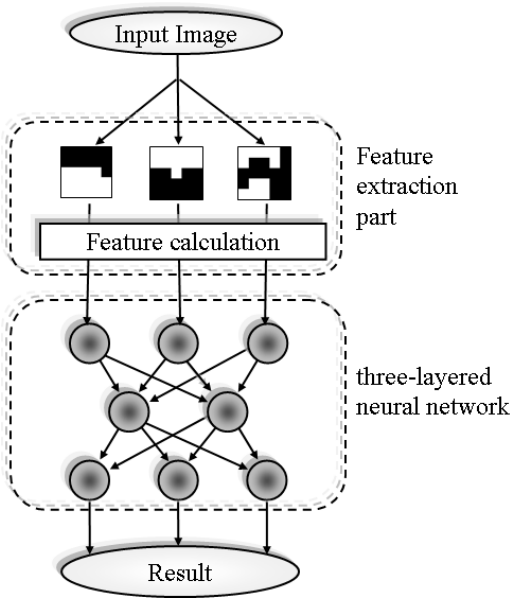

Fig. 1: Contriving recognition model.

Table 1: Parameter for GA

Parameter	
Size of dot pattern	5×5
Number of dot pattern	10
Population size	100
Generation number	15000
Selection type	roulette + elite preservation
Crossover rate	0.8
Mutation rate	0.03
Crossover type	uniform (regard size of dot pattern)

Table 2: Parameter for Neural Network

Parameter	
Size of dot pattern	5×5
Number of dot pattern	10
Input layer	10
Hidden layer	6
Output layer	5

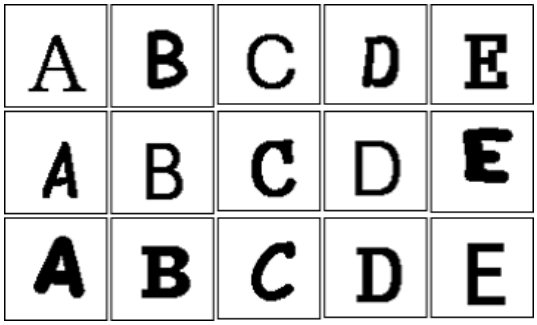

Fig. 2: Sample of images-‘A’,‘B’,‘C’,‘D’ and ‘E’.

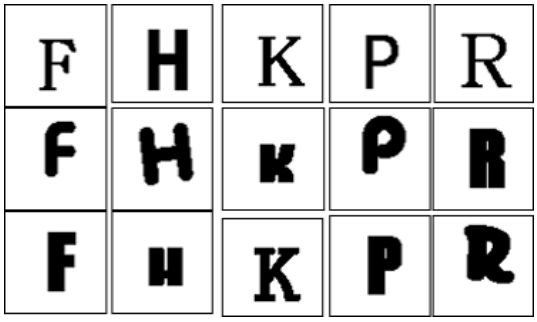

Fig. 3: Sample of images-‘F’,‘H’,‘K’,‘P’ and ‘R’.

2.3. Conventional experiment

We verifies the effectivity of the recognition model which contrived by conventional method experiment. The conventional experiment is (a), (b), (c) and (d).
(a) Conventional experiment 1 : The neural network recognition of the conventional method.
This is recognized by the neural network by using the feature quantity of the pixel.
(b) Conventional experiment 2 : The template matching.
The template matching seeks the similarity of the beforehand decided pattern (template) and the target-pattern.
(c) Conventional experiment 3 : The recognition which used a simple dot pattern.
We create an assumed simple dot pattern and do recognition by the dot pattern.
(d) Conventional experiment 4 : The recognition which used a modeling dot pattern.
We create the dot pattern which resembles an original image and do recognition by the dot pattern.

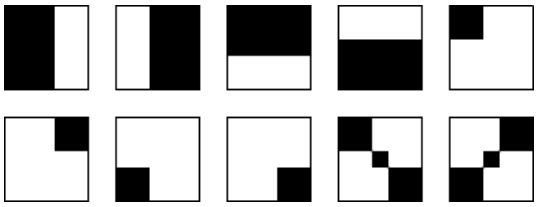

Fig. 4: Simple dot pattern.

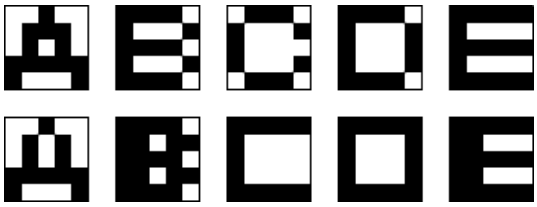

Fig. 5: Modeling dot pattern-‘A’,‘B’,‘C’,‘D’and ‘E’.

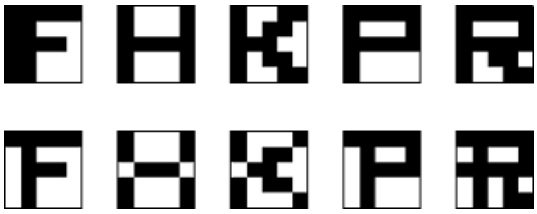

Fig. 6: Modeling dot pattern-‘F’,‘H’,‘K’,‘P’and ‘R’.

3. EXPERIMENT SETTIN

In this section, the way of several experiments to use a multifont alphabet image is described. The data set (experimental data 1, experimental data 2) to have used for the experiment below is shown. Experiment data 1 :‘A’,‘B’,‘C’,‘D’and ‘E’. Experiment data 2 : ‘F’,‘H’,‘K’,‘P’and ‘R’. We created an image from 88 kinds of fonts, chose 25 images as the learning set (125 images) from its inside and chose the others as the unknown image set (315 images). A sample image is shown in Figure 2 and Figure 3.

Next, the setting of a conventional experiment is described. The image size of the template matching and the neural network recognition is 20 $\times$ 20 pixel.

Moreover, the dot size of the Simple dot pattern and the Modeling dot pattern is 5 $\times$ 5 pixel and the number of the dot patterns is 10. Created Simple dot pattern is shown in Figure 4. Also, created Modeling dot pattern is shown in Figure 5 and Figure 6.

4. EXPERIMENTAL RESULTS

This chapter describes the experimental result which used a learning image and a unknown image. The dot pattern which was generated by experimental data 1 is shown in Figure 7 and the dot pattern which was generated by experimental data 2 is shown in Figure 8.

Moreover, the result of experimental data 1 is shown in Table 3 and the result of experimental data 2 is shown in Table 4. The recognition model which contrived from the result in Table 3 and Table 4 showed a high correct answer rate compared with the conventional experiment. From this result, we proved that the contriving recognition model was

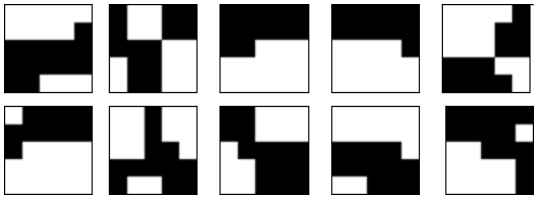

Fig. 7: Obtained dot patterns using GA (multifont ‘A’,‘B’,‘C’,‘D’ and ‘E’).

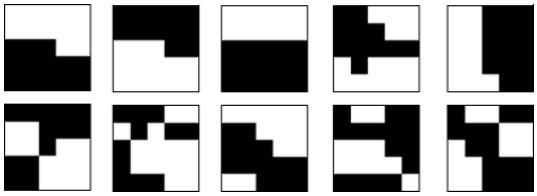

Fig. 8: Obtained dot patterns using GA (multifont ‘F’,‘H’,‘K’,‘P’and ‘R’).

Table 3: Comparison of classification accuracy rate-‘A’,‘B’,‘C’,‘D’ and ‘E’.

	learning images	unknown images
Contriving recognition model	100%	95.6%
Template matching	81.6%	78.7%
Neural networ	100%	84.4%
Simple dot pattern	100%	92.7%
Modeling dot pattern	100%	93.9%

Table 4: Comparison of classification accuracy rate-‘F’,‘H’,‘K’,‘P’ and ‘R’.

	learning images	unknown images
Contriving recognition model	100%	94.6%
Template matching	64.0%	73.0%
Neural networ	100%	73.0%
Simple dot pattern	100%	92.7%
Modeling dot pattern	100%	92.1%

the recognition model which is powerful in the transformation of the shape.

5. CONCLUSIONS

In this paper, we contrived the recognition model which got an idea from the Figure Alphabet Hypothesis. The composition of the model combined the feature extraction part and the neural network which has a dot pattern. Then, it experimented using the unknown image. As a result, as for the experimental result of the contriving recognition model, the correct answer rate which is higher than a conventional way was gotten. Therefore, we could prove the effectivity of the contriving recognition model.

The contriving recognition model has the characteristic which is shown next.
(1) Because the transformation can be absorbed by the simple dot pattern, it is the recognition which is powerful in the transformation of the image.
(2) Because the size and the number of the dot pattern are small, the learning of a neural network and the operation time of the recognition become short.
(3) Because the size and the number of the dot pattern are small, it is possible to do a circuit scale small.
We’ll plan the research a recognition of natural image by the contriving recognition model in the future.

6. REFERENCES

[1] Biederman, I. (1987). “Recognition-by-Components: A Theory of Human Image nderstanding” Psychological Review, 94, 115-147.

[2] Biederman, I., Cooper, E. E. (1991). “Priming contour-deleted images: Evidence for intermediate representations in visual object recognition” Cognitive Psychology, 23, 393-419. ain.

[3] Yoshiyama, K., Uka, T., Tanaka, H. and Fujita, I. “Architecture of binocular disparity processing in monkey inferior temporal cortex” Neurosci. Res. 48:155-167, 2004.

[4] Fujita, I. “The inferior temporal cortex : Architecture, computation, and representation” J. Nuo-cytol., 31 : 359-371, 2002. e r.

[5] DeFelipe, J., Elston, G. N., Fujita, I, et al. “Neocortical circuits : Evolutionary aspects and specificity versus non-specificity of synaptic connections” Remarks, main conclusions and general comments and discussion. J. Neurocytol., 31 :387-416, 2002.

[6] Fujita, I., Tanaka, K., Ito, M., Cheng, K. “Columns for visual features of objects in monkey inferotemporal cortex” Nature, 360: 343-346, 1992.

[7] Kenji Saiki, Ryoji Ohira, Tomoharu Nagao. “FINDING OPTIMIZED OBJECT ALPHABET USING GA” IWAIT2008, 2008.