Associative Learning Method in Hyper-Column Model

Shimada, Atsushi
Department of Intelligent Systems, Kyushu University

Tsuruta, Naoyuki
Department of Electronics Engineering and Computer Science, Fukuoka University

Taniguchi, Rin-ichiro
Department of Intelligent Systems, Kyushu University

出版情報：Proceedings of the Eleventh International Symposium on Artificial Life and Robotics, 2006-01

Proceedings of the 11th International Symposium on
Artificial Life and Robotics, AROB 11th '06
Beppu, Oita, JAPAN, Jan. 2006

Associative Learning Method in Hyper-Column Model

Atsushi Shimada
Department of Intelligent Systems
Kyushu University
6-1 Kasuga-koen Kasuga Fukuoka 816-8580

Naoyuki Tsuruta
Department of Electronics Engineering
and Computer Science
Fukuoka University
8-19-1, Nanakuma, Johnan, Fukuoka 814-0180

Rin-ichiro Taniguchi
Department of Intelligent Systems
Kyushu University
6-1 Kasuga-koen Kasuga Fukuoka 816-8580

Abstract

In this paper, we propose an associative learning method in Hyper-Column Model (HCM). HCM is a model to recognize images, and consists of Hierarchical Self-Organizing Maps (HSOM) and Neocognitron (NC). HCM complements disadvantages of HSOM and NC, and inherits advantages from them. There is a problem, however, that HCM does not suit general image recognition in HCM since its learning method is an unsupervised one with competitive learning which is used by Self-Organizing Map (SOM). Therefore, we extended HCM to a supervised learnable model with an associative learning of SOM. We have found that an ability of HCM with the associative learning is superior to the one with unsupervised learning.

1 Introduction

Recently, applications of Artificial Neural Networks (ANNs) have expanded into general image recognition such as face recognition and visual surveillance. In these applications, there are three problems.

1. In spite of the high dimension of the input images, the regions of object parts occupy small areas of the whole space. Therefore, it is necessary to reduce the dimension to eliminate the redundancy and to allow the network to learn according to the region actually expanding the object.
2. The absolute dimensionality of the image region is very large even if the redundancy is eliminated, since images have large variations as a function of object locale, illuminant, and so on. Therefore, techniques to reduce the system size to a realizable scale are necessary.
3. In such cases, ANNs need a very large number of neurons. Therefore, learning methods which depend on the initial states of the connection weights cannot perform well, even if a network model such as the multilayered perceptron is theoretically very powerful. Thus, learning methods whose performance does not depend on the initial states of connection weights are needed.

Tsuruta proposed Hyper-Column Model (HCM)[1] which is a new image recognition model combining Hierarchical Self-Organizing Map (HSOM)[2] and Neocognitron (NC)[3]. The learning method is quite simple with unsupervised learning, but it is powerful enough since it does not depend on the initial states of the connections. In addition, HCM can reduce the dimensionality of general images and can perform better than HSOM and NC, since HCM overcomes the disadvantages of HSOM and NC and inherits their advantages directly. HCM cannot, however, show high ability when the boundary between categories is very complicated, such a case that objects of different categories have similar features. Therefore, a new learning algorithm for the HCM is proposed. The proposed algorithm is associative learning method, which is very powerful to separate the similar input data which should belong to different categories on the feature map.

2 HSOM and NC

2.1 Hierarchical Self-Organizing Map (HSOM)

The HSOM is a two layers of SOM network connected as any feedforward neural network: every unit in the sending layer is connected to every unit in the receiving layer. The basic SOM defines a mapping from the input data space onto a regular array of neurons. All neurons $(1 \cdots u \cdots N)$ are fed the same input data $\boldsymbol{I}$. Each neuron has a weight vector $\boldsymbol{W}_u$. When the neurons are fed $\boldsymbol{I}$, they are activated according to the similarity between $\boldsymbol{I}$ and $\boldsymbol{W}_u$. In practical applications the Euclidian distance is usually used as the similarity measure, only neuron c can be activated, where

$$||\boldsymbol{I} - \boldsymbol{W}_c|| = \min_u(||\boldsymbol{I} - \boldsymbol{W}_u||) \tag{1}$$

In the training phase, each time a training data item is input, te winner is selected according to Eq. (1) and is trained according to the following equations:

$$\boldsymbol{W}_u(t+1) = \boldsymbol{W}_u(t) + h_{cu}(\boldsymbol{I}(t) - \boldsymbol{W}_u(t)) \tag{2}$$

$$h_{cu} = \alpha(t) \cdot \exp\left(-\frac{||r_c - r_u||^2}{2\sigma^2(t)}\right) \tag{3}$$

h_{cu} is a neighborhood kernel function. With increasing $||r_c - r_u||$ and t, h_{cu} converges to zero. $\alpha(t)$ is a monotonically decreasing function of t $(0 < \alpha(t) < 1)$, and $\sigma^2(t)$ defines the width of the kernel.

In the case of HSOM, the neurons $(1 \cdots v \cdots M)$ in the second layer are fed the index of the winner neuron in the first layer as input data. The training algorithm in the second layer is same as that in the first layer but the number of neurons M is smaller than N. This small number of neurons help the HSOM to integrate the features extracted in the first layer.

Characteristics of the HSOM are summarized in the following points.

1. HSOM solves the complex region problem, when a sufficiently dense data set is given for continuous variation of the input data.

2. HSOM cab be a good preprocess to resolve the dimensional reduction problem.

In case where the HSOM is used directly for general image recognition, the following three problems, however, arise.

1. Image recognition methods based on HSOM are regarded as a memory-based method. Therefore, in case where the dimension and the size of data distribution are large, the network size increases.

2. Maps are organized according to the continuity of data distribution in the space defined by the distance between images using template matching. Therefore, it is hard to organize maps in cases where the distance between images does not smoothly vary with locale and scale of target object. In such cases, a heuristic, such as image blurry, must be adapted or the number of training samples must be increased so that the variations of distance are smooth enough.

3. The recognition method is also based on the distance of images and, therefore, is regarded as one of nearest neighbor method. Therefore, a high accurate segmentation of target region and a normalization of locale and size should be needed.

2.2 Neocognitron (NC)

Neocognitron (NC) is proposed by Fukushima[3] as a hierarchical network consisting of several layers of neuron-like cells. The lowest stage of the network is the input layer. Each succeeding stage has a layer consisting of cells called S-cells followed by another layer of cells called C-cells. S-cells are the feature-extracting cells. The C-cells are inserted on the network to allow for positional errors in the features.

The "structural" advantages of NC are summarized in the following two points, which can alleviate the disadvantages of HSOM.

1. Every feature map is rather small since NC has a hierarchical structure of divide-and-conquer type. This characteristic resolves the disadvantage 1 of HSOM.

2. Toward the disadvantage 2 and 3 of HSOM, NC does not need any preprocesses of blurring nor normalization due to its shift invariant recognition mechanism. NC, also, needs small number of training sample data because of its learning mechanism.

The original NC has the following two disadvantages when it is directly applied to general image recognition. The advantages of HSOM match with those disadvantages.

1. The learning process of NC is strongly depends on the initial state of weight vectors, and the feature extraction does not often perform well due to its simple competitive learning without neighborhood learning.

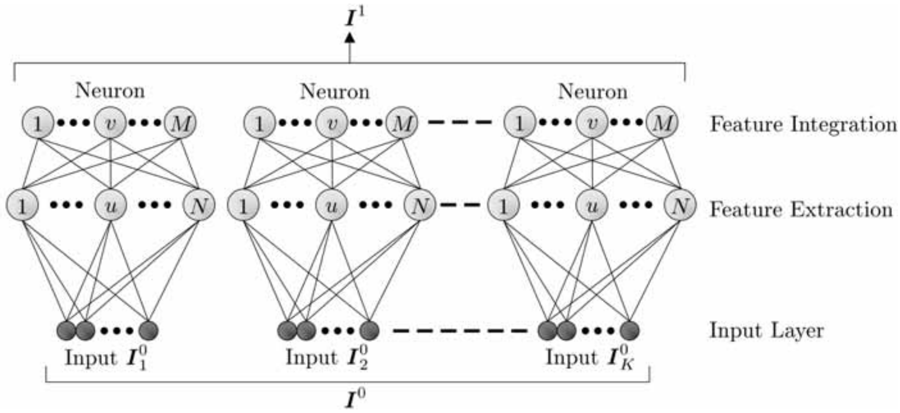


Figure 1: Structure of HCM

2. The shift integration layer integrates only shifted patterns, and does not requantize feature spaces. Therefore, its dimension reduction is not enough for variation of target shape, which requires large number of network neurons.

3 Hyper-Column Model (HCM)

Hyper-Column Model (HCM) is proposed for visual recognition of objects with variations in its position, size, and orientation[1]. HCM is a kind of competitive neural networks with unsupervised learning. The network is composed of hierarchical layers derived from NC by replacing the unit cell plains (each C-cells and the lower directly connecting S-cells) in NC with two-layers HSOM. These HSOM cell plains allow features extraction through the first SOM layer followed by features integration in the second SOM layer, as shown in Fig. 1. This feature integration process allows the cell plains to integrate more features in lower number of neurons. NC does not perform such feature integration process.

3.1 Associative Learning in HCM

The structure of HCM is similar to the one of NC, but HCM uses unsupervised learning algorithm that is used in SOM. In general image recognition problems, it is known that the recognition ability with supervised learning method is superior to the one with unsupervised learning method. Therefore, introducing a supervised learning method into HCM is needed.

We introduce a method of supervised learning in SOM proposed by Ichiki[4] into HCM. In this learning method, we can regard the input vector as the one which is composed of different two components; an input part $\boldsymbol{I}$ and an associative part $\boldsymbol{T}$. As a result, the networks can be considered as a supervised learning machines.

$$\boldsymbol{X} = a \begin{pmatrix} \boldsymbol{I} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \boldsymbol{T} \end{pmatrix} \tag{4}$$

where $a > 1$ in order that the input part can affect the information of the map more than the output part does. The weight vector $\boldsymbol{W}_u$ is represented as follows,

$$\boldsymbol{W_u} = \begin{pmatrix} \boldsymbol{W}_u^I \\ \boldsymbol{W}_u^T \end{pmatrix} \tag{5}$$

4 Experimental Results

4.1 Condition

The input data were images of human hands which consist of 10 categories shown in Fig. 2. Each category has 10 images for training data, and 500 images for test data. The size of images was 100×100, and each pixel had an 8-bit gray value. The neurons in each feature extraction layer were arranged in a ring shape. In experiment, training data were learned by the traditional learning (HCM) and by the associative learning (AHCM).

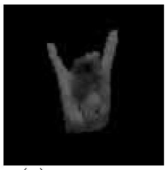
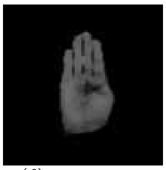
4.2 Learning result

The maps generated by each HCM are shown in Fig. 3. The horizontal axis shows the sample number. From number 0, each 10 samples belong to the same category. The vertical axis shows the neuron number. HCM cannot generate one neuron cluster for the whole of one category 1, 2, 6, 7 and 8. On the other hand, AHCM can do it in most cases.

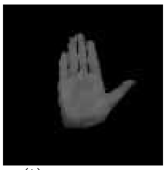

(a) category 1

(b) category 2

(c) category 3

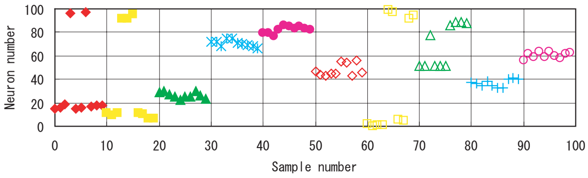
(d) category 4

(e) category 5

(f) category 6

(g) category 7

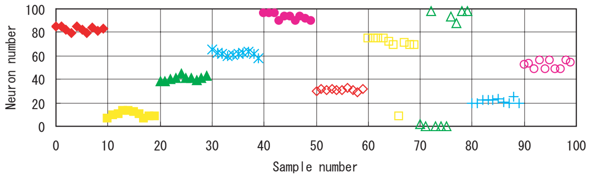
(h) category 8

(i) category 9

(j) category 10

Figure 2: Training images



(a) Traditional learning



(b) Associative learning

Figure 3: Correspondence between maps of HCM and sample data in experiment

4.3 Recognition accuracy

The average rate of correct recognition of HCM was 72.5%. The images of category 7 tended to be misunderstood as category 2. There are some overlaps between the neurons which belong to category 2 and category 7 as can be expected from Fig. 3(a). The images which belong to category 9, also, tended to be misunderstood as category 6. In most images which were not recognized correctly, the position of thumb was nearby the palm. We think that HCM with unsupervised learning could not create an appropriate boundary between category 6 and category 9 since the images of category 9 are similar to the ones of category 6.

On the other hand, the average rate of correct recognition of AHCM was 93.3%. The associative learning method gave good results without decreasing the recognition rates.

5 Conclusions

An associative learning method in Hyper-Column Model (HCM) was proposed. In the experiments described in this paper, the recognition ability of HCM with the associative learning was superior to the traditional learning. However, HCM has a problem that it runs out of neuron when the number of category increases. We are now researching for reconstruction of the network and incremental learning.

References

[1] Naoyuki Tsuruta, Rin-ichiro Taniguchi, and Makoto Amamiya. "Hypercolumn Model: A Combination Model of Hierarchical Self-Organizing Maps and Neocognitron for Image Recognition". *Systems and Computers in Japan*, 31(2):49–61, 2000.

[2] J. Lampinen and E. Oja. "Clustering properties of hierarchical self-organizing maps". *Journal of Mathematical Imaging and Vision*, 2(3):261–272, 1992.

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

[4] Ichikin H., Hagiwara M, and Nakagawa M. "Kohonen feature maps as a supervised learning machine". *Proc. of International Conference on Neural Networks*, pages 1944–1948, 1993.