Object Recognition Using a Self-Organizing Map for an Autonomous Mobile Robot

Masayoshi Tabuse **Haruna Kaneko**
Department of Environmental Information, Kyoto Prefectural University, Kyoto, 606-8522, Japan
tabuse@kpu.ac.jp

Abstract

In this research, we propose a control system by using the self-organizing map, a neural network and the genetic algorithm. The control system autonomously extracts visual information and decides behaviors. We investigate that a mobile robot controlled by this system performs a given task very well.

Key words: object recognition, Self-organizing map, neural network, mobile robot

1 Introduction

Recently, many people have been researching autonomous robots extensively. One purpose of this research is to build the robots that are able to behave in unknown or dynamically changing environments. In the case of a robot with visual sensors, most of methods are that useful information is extracted using image processing and robot behaviors are decided based on this information. However, appropriate image processing methods are generally constructed by a designer who has a good knowledge of tasks, so that autonomy and flexibility of a robot may be reduced [1]. Therefore we propose a control system in which a robot autonomously extracts visual information and decides behaviors.

In this system, we use a CCD camera as a visual sensor. Images obtained from the camera inputs Kohonen's self-organizing map [2], which classifies images, and a robot recognizes objects in the environment. A behavior decision part consists of a neural network and parameters of the neural network are chosen so that the robot accomplishes a given task by using a genetic algorithm [3]. We examine this system using a miniature robot Khepera with a CCD camera. A task is that the robot moves in a field, finds a coloring block and carries it to a same coloring area. We investigate that Kohonen's self-organizing map classifies a floor, a wall, coloring blocks and coloring areas and that the robot accomplishes the task very well.

2 Outline of the System

2.1 A mobile Robot

We use a miniature mobile robot Khepera II in our experiments as shown in Figure 1. This robot is 70mm in diameter and has two independent wheels and eight infrared proximity sensors (six in front and two in rear). We attach a gripper to this robot so that it grips a small block. Furthermore, we place a small CCD camera (KEYENCE CK-200) on the gripper to capture images.

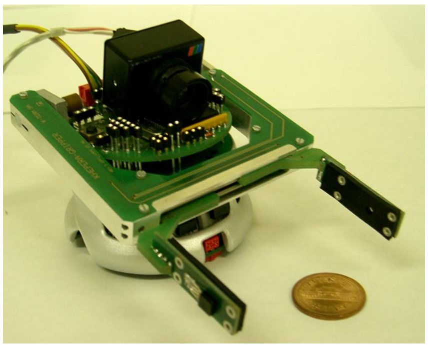

Figure 1: A mobile robot Khepera II

2.2 Control System

The control system consists of Kohonen's self-organizing map (SOM) and a two-layer neural network (NN). Kohonen's self-organizing map classifies images and the neural network control each wheel and a gripper. A CCD camera mounted on Khepera captures front images. These images are inputted into SOM. The output data from SOM are fed to a two-layer neural network. Weights of the neural network are determined using the genetic algorithms so that

Khepera performs appropriate behaviors. The outline of control system is shown in Figure 2.

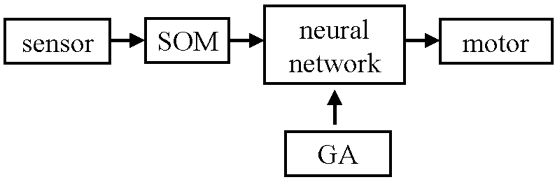

Figure 2.: The outline of control system

2.3 SOM

The self-organizing map (SOM) is a neural network of unsupervised learning proposed by T. Kohonen in the early 1980s. The goal is to discover some underlying structure of the data. SOM consists of two layer structures. One is an input layer and the other is an output layer, which is usually represented by 2 dimensional grid. The network is fully connected, i.e., all nodes in input layer are connected to all nodes in output layer, as shown in Figure 3.

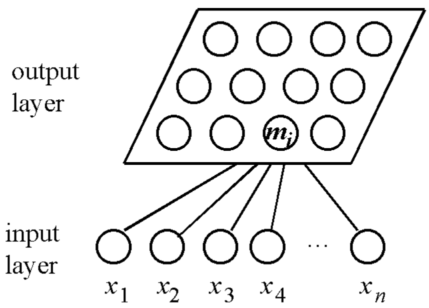

Figure 3: SOM

The learning algorithm for *n* dimensional input vectors is as follows:

(1) Properly choose a reference vector of i-th output node

$$\vec{m}_i = (m_{i1}, m_{i2}, \dots, m_{in}),\ i = 1,2,\dots,M ,$$

connected with *n* input nodes, where *M* is the number of output nodes.

(2) An input vector $\vec{x} = (x_1, x_2, \dots, x_n)$ is compared with all the reference vectors $\vec{m}_i$. The best-matching output node where the reference vector is most similar to the input vector in some metric (e.g. Euclidean) is identified. This best matching node is called the winner. The winner node *c* is determined by

$$c = \arg\min_i \| \vec{x} - \vec{m}_i \|.$$

(3) The reference vectors of the winner and its neighboring nodes on the grid are changed towards the input vector according to the following equation:

$$\vec{m}_i(t+1) = \vec{m}_i(t) + h_{ci}(t)[\vec{x}(t) - \vec{m}_i(t)],$$

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(-\frac{\| \vec{r}_c - \vec{r}_i \|^2}{2\sigma^2(t)}\right),$$

where t=0,1,2,… is a discrete learning time, $h_{ci}(t)$ is a neighborhood function, $\alpha(t)$ is a learning rate, $\sigma(t)$ is a decreasing function which defines a radius of neighborhood and $\vec{r}_c$ and $\vec{r}_i$ are coordinate vectors of the winner node *c* and the neighborhood node *i* on the grid, respectively.

(4) Repeat processes (2) and (3).

According to this algorithm, the network organizes itself and a self-organizing map for input vectors is built. This map classifies input vectors. In Figure 5 we show a self-organizing map obtained in our experiments. Light and shade of the map represent similarity of reference vectors. Light color means that reference vectors of both sizes are similar. Thus we guess that a robot with our control system recognizes objects without appropriate image processing methods constructed by a designer.

2.4 Neural network and genetic algorithm

A two-layer neural network controls left and right wheels and a gripper according to output data of SOM as input data. We think that a neural network is good controller because it outputs continuous signals for various input data. Usually, learning methods of the neural network are supervised learning. In supervised learning, we must have knowledge of correct outputs of the neural network for any input data. However, in dynamical changing environments, we suppose that it is impossible to teach correct outputs in any cases. Therefore, We use the genetic algorithm to find appropriate weights of the network to perform a given task.

3 Experiments

3.1 Experimental Environment

Figure 3 shows the environment in our experiments. The color of a floor and a wall is white and we set red, blue and green blocks and the same coloring areas on the floor. A task is that Khepera carries a block to the same coloring area.

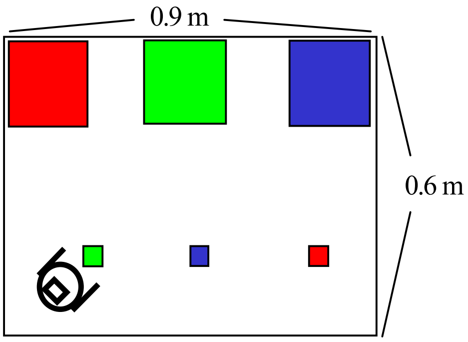

Figure 3: Experimental environment

3.2 Learning methods

First we capture total 190 sample images of a block, a coloring area, a wall and a floor using a CCD camera mounted on Khepera. Each image is 320×240 pixels, but to reduce the dimension of input data and to obtain a image of the nearest object, we use only the lower half of the image after dividing into tiles (tile size : 20×20) and averaging in each tile. A capture image and its lower half average image are shown in Figure 4. Then 16×6 = 96 dimensional images are the input data of SOM. SOM consists of 96 nodes in input layer and 6×8 nodes in output layer, and SOM classifies input data using the learning algorithm. Figure 5 shows a self-organizing map for 190 sample images. Images of a coloring block, a floor or a wall and a red area are classified in the left upper side, in the center and in the lower side of this map, respectively. The number on the map represents a sample image number.

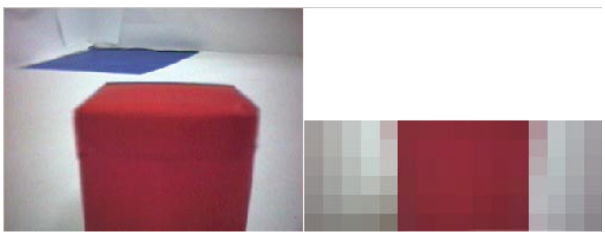

Figure 4: A example of a capture image and its lower half average image

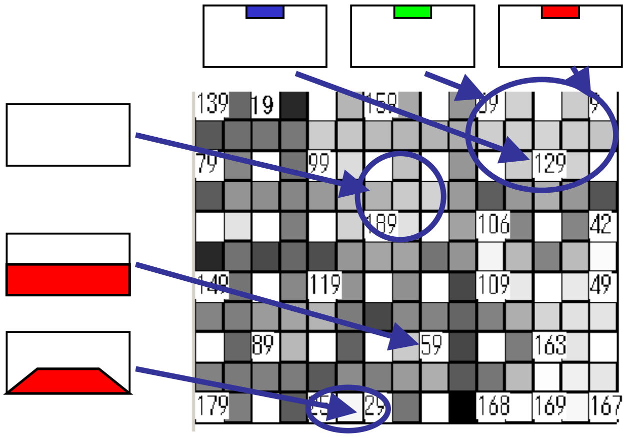

Figure 5: SOM of our experiments

Next we determine weights of the neural network to perform a given task using the genetic algorithm in computer simulation. The neural network consists of two layers, as shown in Figure 6. Input data are 48 dimensional data from SOM and 3 internal states of Khepera. Internal states denote that Khepera is gripping a block, i.e.,

$$\text{internal state} = \begin{cases} (0,0,0) & \text{for gripping no block} \\ (1,0,0) & \text{for gripping a red block} \\ (0,1,0) & \text{for gripping a green block} \\ (0,0,1) & \text{for gripping a blue block} \end{cases}$$

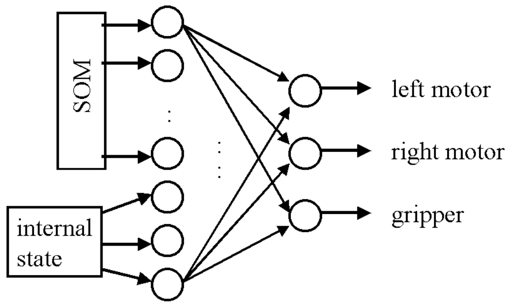

Figure 6: The neural network

In this research, we investigate object recognition by using SOM. Therefore, we consider that learning behaviors of Khepera are finding a block, approaching a block, carrying a block to a same coloring area and outputting signals of gripping and placing a block to the gripper. The behaviors of gripping and placing a block are constructed by hand so that if the signal to the gripper indicates gripping a block and infrared sensors detect a object in front, Khepera grips a block and that if the signal indicates placing a block, Khepera places a block.

In the genetic algorithm we set that the population has 50 individuals, it runs for 2000 generations and individuals run 200 steps for each generation. Each individual has a chromosome, which represents weight of the neural network in a real number. At each generation, individuals with high fitness are selected and produce descendants using crossovers and mutations of their chromosomes. The fitness function is given by

$$fitness = \begin{cases} 1000 & \text{for gripping a block} \\ 1000 & \text{for placing a block} \\ & \text{on the same coloring area} \\ -500 & \text{for going back or rotating} \\ & \text{in a free space} \\ 10 & \text{for capturing a block} \\ -100 & \text{for missing a block} \end{cases}$$

After 2000 generations, we select an individual with highest fitness, decide weights of the neural network and investigate Khepera's behaviors with the control system.

3.3 Result

The mobile robot Khepera moves around in the field and finds a coloring block. Then the robot approaches the block, and grips and carries it and places it on a same coloring area. Thus we find that Kohonen's self-organizing map classifies a floor, a wall, coloring blocks and coloring areas and that the robot accomplishes the task very well. Figure 7 and 8 show scenes of gripping and placing a block, respectively.

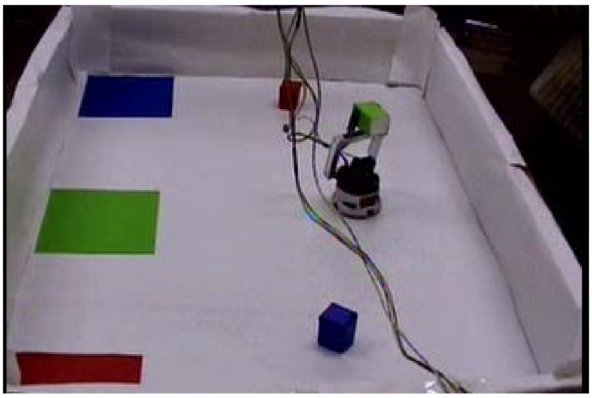

Figure 7: Griping a block

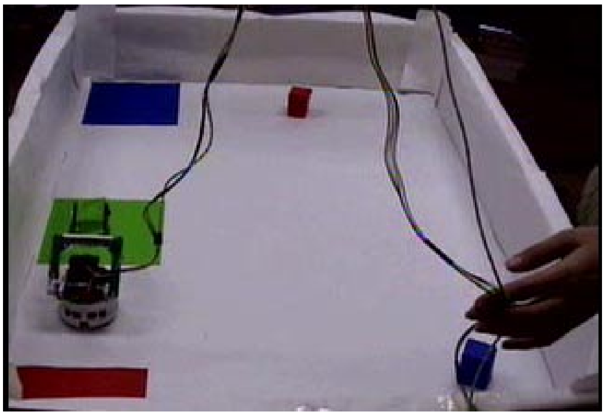

Figure 8: Placement of a block on the same color area

4 Conclusion

We have presented the control system for an autonomous mobile robot using a combination of Kohonen's self-organizing map and neural network. In this system, the robot autonomously extracts visual information and recognizes objects in the environment. Thus the robot finds a colored block, grips and carries it and places it on a same colored area.

In our experiments, we use a map of 6×8 in SOM. If the size of a map is larger, SOM may classify visual images more precisely. However, for a larger size of a map it requires more time to find proper weights of neural network, so that we will find a proper size of a map.

As the future works, we will consider a control system in which a robot recognizes a shape of objects and construct a robust system in dynamically changing environments.

References

[1] K. Shibata, M. Iida, "Acquisition of box pushing by direct-vision-based reinforcement learning", SICE 2003 Annual Conference, pp.2322- 2327, 2003.

[2] T. Kohonen, "Self-Organizing Maps", Springer-Verlag, 1995.

[3] S. Nolfi, D. Floreano, "Evolutionary Robotics; The Biology, Intelligence, and Technology of Self-Organizing Machines", MIT Press, 2000.