The Team Description of Osaka University "Trackies-99"

Sho'ji Suzuki[1], Tatsunori Kato[1], Hiroshi Ishizuka[1],
Hiroyoshi Kawanishi[1], Takashi Tamura[1], Masakazu Yanase[1],
Yasutake Takahashi[1], Eiji Uchibe[1], and Minoru Asada[1]

Dept. of Adaptive Machine Systems, Graduate School of Engineering,
Osaka University, Suita, Osaka 565-0871, Japan

Abstract. This is the team description of Osaka University "Trackies" for RoboCup-99. We have worked two issues for our new team. First, we have changed our robot system from a remote controlled vehicle to a self-contained robot. The other, we have proposed a new learning method based on a Q-learning method so that a real robot can aquire a bhevior by reinforcement learning.

1 Introduction

We are interesting in how a robot acquires a behavior in dynamic environments and how robots cooperate without explicit communication. in the context of cooperative distributed vision [1]. We have applied a Q-learning method, one of major method of reinforcement learning, to real robots and tested in RoboCup competitions[2][3].

However, the performance of the behavior is not enough because;

1. the controll of the robot is not reliable.
2. the applied method is not enough to adapt to the real robot.

In RoboCup-99, we will improve these problems by building a new platform and propsing a new learning method. In the rest of this paper, we describe our new robot and propse a new learning method.

2 The Robot of Osaka University "Trackies-99"

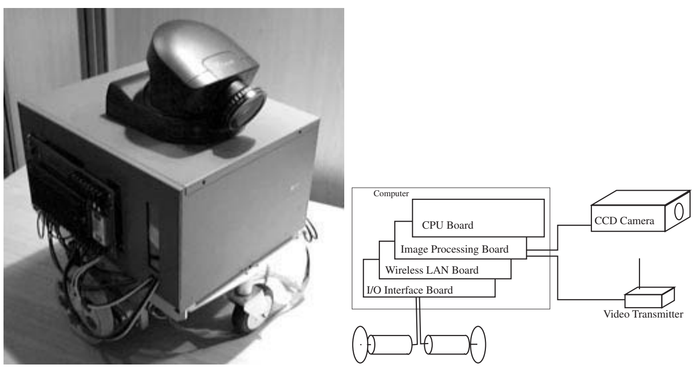
In RoboCup-97 and 98, we used a radio controlled model car as a robot body and equiped a CCD camera and a video transmitter on it. The image captured by the camera was transmitted to a host computer and it sent control signals to the robot. We could test various image processor and software development tools, however, the control of the robot was not reliable due to noises on radio links[2][3] and it brought poor performance of the robot's behavior.

To escape from noise problem, we have build a self-contained robot where the host computer is equiped on the robot. Figure 1(a) shows the robot of the team of Osaka University "Trackies-99". Figure 1(b) shows configuration of its controller including following devices:

M. Veloso, E. Pagello, and H. Kitano (Eds.): RoboCup-99, LNAI 1856, pp. 750–753, 2000.

base of the vehicle is a product of Mechatro Systems. Left and right wheel is driven by a DC motor and other two wheels are caster.
CCD Camera is Sony EVI-D30. Pan and tilt angle of the camera is controllable from CPU via serial communication.
CPU and Operating System Linux is running on Pentium MMX 233MHz.
image processor is Hitachi IP-5005. It is a fast color image processing board where basic operations are installed.
LAN WaveLAN is used for monitoring and debug.
video transmitter is used for monitoring a processed image

The size of the robot is 400[mm] of length, 360[mm] of width, and 450[mm] of height. The weight is 10[Kg].



(a) overview (b) controller

Fig. 1. Robots of Osaka University “Trackies-99”

3 A Learning method for a real robot

We propose a continuous valued Q-learning for real robot applications. Unlike the conventional real-valued Q-learning methods, the proposed method dose not need well-defined quantized state and action spaces to converge. The basic idea for continuous value representation of state, action, and reward in $Q-learning$ is to describe them as contribution vectors of representative states, actions, and rewards.

First, we tessellate the state space into n-dimensional hyper cubes[1] . The vertices of all hyper cubes can be the representative state vectors $\boldsymbol{x}^i = (x_1^i, x_2^i, \cdots x_n^i)$

[1] the unit length is determined by normalizing the length of each axis appropriately

$i = 1, \cdots, N$ (here, N denotes the number of the vertices), and we call each vertex the representative state s_i. The contribution value $w_i^{\boldsymbol{x}}$ for each representative state s_i when the robot perceives the input $\boldsymbol{x} = (x_1, x_2, \cdots x_n)$ is defined as follows:

1. Specify a hyper cube including the input $\boldsymbol{x} = (x_1, x_2, \cdots x_n)$.
2. Tessellate the cube into 2^n hyper boxes based on the input $\boldsymbol{x}$ (see Figure 2 for the two dimensional case)
3. Calculate the volume of each hyper box.
4. Assign the volume $w_i^{\boldsymbol{x}}$ of the box diagonal to the state s_i.
5. If the input $\boldsymbol{x}$ is on the surface of the hyper cube, the volume can be reduced to the area or the length.
6. **Any other contribution** values for the states which do not compose the above cube are all zeros.

Mathematical formulation of the above process is given by

$$w_i^{\boldsymbol{x}} = \prod_{k=1}^{n} l_i(x_k), \tag{1}$$

where

$$l_i(x_k) = \begin{cases} 1 - |x_k^i - x_k| & (\text{if } |x_k^i - x_k| \leq 1) \\ 0 & (\text{else}) \end{cases} \tag{2}$$

Figure 2 shows the case of two-dimensional sensor space. The area $w_i^{\boldsymbol{x}}$ is assigned as a contribution value for state s_i. The summation of contribution values $w_i^{\boldsymbol{x}}$ for the input $\boldsymbol{x}$ is one, that is,

$$\sum_{i=1}^{N} w_i^{\boldsymbol{x}} = 1 \tag{3}$$

Thus, the state representation corresponding to the input $\boldsymbol{x}$ is given by a state contribution vector $\boldsymbol{w^x} = (w_1^{\boldsymbol{x}}, \cdots, w_N^{\boldsymbol{x}})$. Similarly, the action representation corresponding to the output $\boldsymbol{u}$ is given by an action contribution vector $\boldsymbol{w^u} = (w_1^{\boldsymbol{u}}, \cdots, w_M^{\boldsymbol{u}})$, where M denotes the number of the representative actions a_j in the tessellated action space.

To show the validity of the method, we applied the method to a vision-guided mobile robot of which task is to chase the ball. Figure 3 shows a sequence of the aquired behavior. Although the task was simple, the performance was quite impressive.

4 conclusions

This work was supported by the Cooperative Distributed Vision project in the Research for the Future Program of the Japan Society for the Promotion of Science (JSPS-RFTF96P00501).

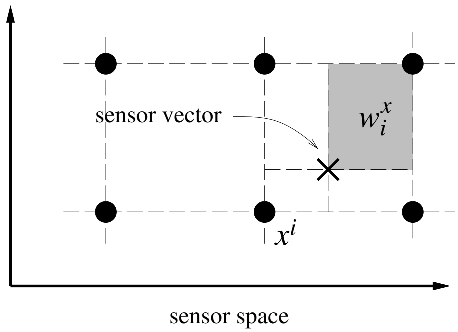


Fig. 2. Calculation of contribution value w_i for the representative state s_i in the case of two-dimensional state space

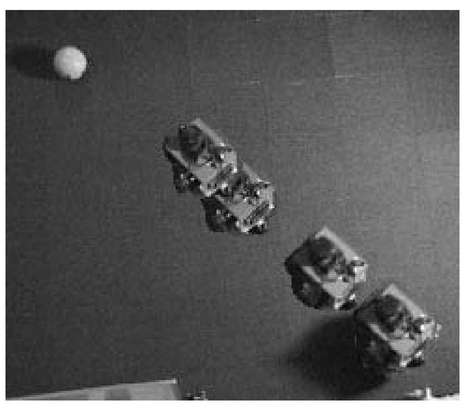


Fig. 3. A part of video image sequence

References

1. T. Matsuyama. Cooperative Distributed Vision – Dynamic Integration of Visual Perception, Action, and Communication –. In *Proc. of Image Understanding Workshop*, 1998.
2. S. Suzuki, Y. Takahashi, E. Uchibe, M. Nakamura, C. Mishima, and M. Asada. "Vision-Based Learning Towards RoboCup: Osaka University 'Trackies' ". *RoboCup-97: Robot Soccer World Cup I*, Springer, pp.305–319, 1997.
3. S. Suzuki, T. Kato, H. Ishizuka, Y. Takahashi, E. Uchibe, and M. Asada. "An Application of Vision-Based Learning in RoboCup for a Real Robot with an Omnidirectional Vision System and the Team Description of Osaka University "Trackies" ", *RoboCup-98: Robot Soccer World Cup II*, Springer, 1999(to be published).