Vision-Based Learning for Real Robot: Towards RoboCup

S. Suzuki, Y. Takahashi, E. Uchibe,
M. Nakamura, C. Mishima, and M. Asada

Dept. of Adaptive Machine Systems, Osaka University, 2-1, Yamadaoka, Suita, Osaka 565, Japan
ssuzuki@robotics.ccm.eng.osaka-u.ac.jp

Abstract

The authors have applied reinforcement learning method to a mobile robot with vision system. We selected a task for the robot from skills for playing soccer. In the first stage, a robot learned to shoot a ball into a goal. In the second stage , we set up an opponent just before the goal, that is, a goal keeper, and make the robot learn to shoot a ball into a goal avoiding the goal keeper. This paper describes several research issues for RoboCup with real robots along with our research projects.

1 Introduction

Building robots that learn to perform a task has been acknowledged as one of the major challenges facing AI and Robotics. Reinforcement learning has recently been receiving increased attention as a method for robot learning with little or no *a priori* knowledge and higher capability of reactive and adaptive behaviors [5]. In the reinforcement learning scheme, a robot and an environment are modelled by two synchronised finite state automatons interacting in discrete time cyclical processes. The robot senses the current state of the environment and selects an action. Based on the state and the action, the environment makes a transition to a new state and generates a reward that is passed back to the robot. Through these interactions, the robot learns a purposive behavior to achieve a given goal.

As a test bed for real robot applications of the reinforcement learning method, we have selected soccer playing robots [8]. In the first stage[2][1], a robot learned to shoot a ball into a goal given the state space in terms of the sizes and the positions of both the ball and the goal in image. In the second stage [4][3][7] , we set up an opponent just before the goal, that is, a goal keeper, and make the robot learn to shoot a ball into a goal avoiding the goal keeper. The behavior of the opponent is scheduled for the learner to efficiently obtain the desired behavior .

In this paper, we explain our robot system and summarise our research issues involved in realizing RoboCup Initiative with real robots.

2 Real Robot System

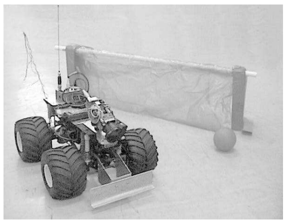
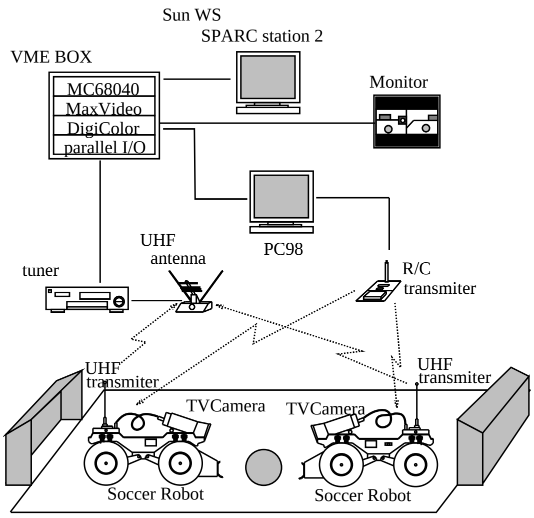


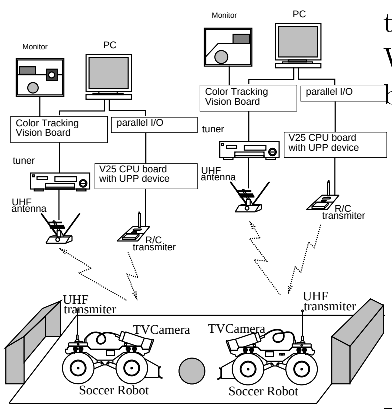
Figure 1: A picture of the radio-controlled vehicle

The environment consists of a ball, two goals, four lines and a keeper robot. The aim of our research is that the robot learn a behavior to accomplish a task without knowing the location and the size of the objects, the size and the weight of the ball, keeper robot, any camera parameters such as focal length and tilt angle, or kinematics/dynamics of itself. A behavior is realised as a sequence of action primitives which is a motion(going straight, turning right, turning left, etc) implemented to the robot initially.

We use a radio controlled car as a robot's body and moves using a 4-wheel steering system. The robot has a single color TV camera and the effects of an action against the environment can be informed to the robot only through the visual information. The robot is shown in **Fig.1**.



(a) VME based system



(b) PC based system

Figure 2: Configuration of robot controller

Fig.2 shows a configuration of the real mobile robot system. The robot is controlled by the remote computer. Now we use VME based system and PC based system. The image taken by a TV camera mounted on the robot is transmitted to a UHF receiver and processed. The remote computer decide the action of the robot and generate a control signal to execute it.

In VME base system, Datacube MaxVideo 200 a real-time pipeline video image processor is used for image processing, and VxWorks OS on MC68040 CPUs which are connected with host Sun workstations via Ether net is used for vehicle control. In order to simplify and speed up the image processing time, we painted the ball, the goal, and the opponent in red, blue, and yellow, respectively. The input NTSC color video signal is first converted into HSV color components in order to make the extraction of the objects easy.

In PC base system, Fujitsu color tracking vision is used for image processing.

3 Applying Q-learning to a real robot

We have applied reinforcement learning method to a mobile robot with vision. In this section, we summarise our issues.

3.1 Shooting Behavior Acquisition[2]

In the first stage, the task for a mobile robot is to shoot a ball into a goal as shown in **Fig.3**. We assume that the environment consists of a ball and a goal.

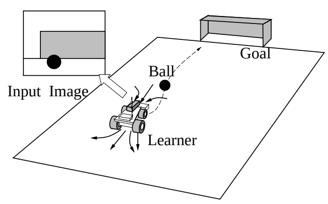


Figure 3: The task is to shoot a ball into a goal

We define sub-states and action space and give to the robot. Sub-states are defined corresponding to the position and the size of the ball or goal which are naturally and coarsely classified image(**Fig.4**). The action space is defined as follows. Each action is regarded as an action primitive. The robot continues to take one action primitive at a time until the current state changes. This sequence of the action primitives is called an action. In the above case, the robot takes a forward motion many times until the state "the goal is far" changes into the state "the goal is medium".

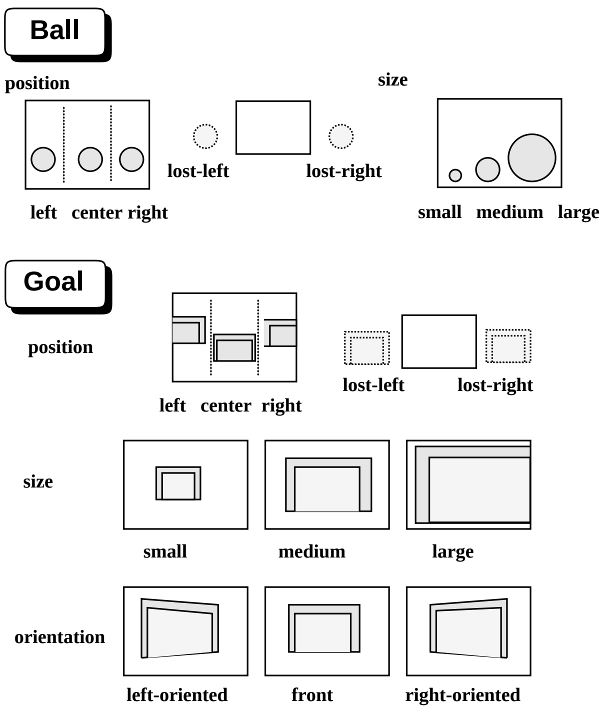


Figure 4: The ball sub-states and the goal substates

3.2 Shooting a Ball with Avoiding an Opponent [3, 4, 7]

In the second stage, we set up an opponent just before the goal, that is, a goal keeper, and make the robot learn to shoot a ball into a goal avoiding the goal keeper (See Fig.5). The basic idea is first to obtain the desired behavior for each subtask, and then to coordinate two learned behaviors. For the first subtask (shooting behavior), we have already obtained the learned policy by using the state space shown in 4. For the second subtask (avoiding behavior), we add the sub-states for the opponent that consist of the size and position of it in image.

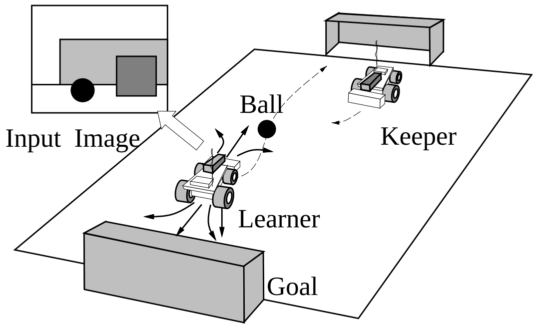


Figure 5: The task is to shoot a ball into the goal avoiding an opponent.

4 Self Construction of the State Space

Two major problems exist in applying reinforcement learning method to real robot tasks: how to construct the state space, and how to reduce the learning time. The authors have proposed another method by which a robot learns purposive behavior within less learning time by incrementally segmenting the sensor space based on the experiences of the robot[6].

For state construction, the following two policies can be considered to segment the state space.

A: Segment the state if the prediction of sensor outputs is incorrect.

B: Segment the state if the same action causes the desirable or undesirable result (ex., transition to the goal states or non-goal states) even though the prediction itself is correct.

the policy **A** is related to the world model construction by coarse mapping between states and actions far from the goal states. While, the policy **B** is related to the the goal oriented segmentation based on the reinforcement signals. As a result fine mapping between states and actions near the goal states is obtained. Therefore, switching of two policy is necessary to learn purposive behavior.

Based on this idea, we use a linear model of the gradient of sensor outputs for segmentation based on the policy **A**. By this method, the robot learns a shoot behavior shown in Fig.3 in sevral hours. Fig. 6 shows the state space after 72 trials.The state space in term of ball size and goal size is indicated when the position of the ball and the goal are center of the screen and the orientation of the goal is frontal. The numbers of acquired states and data are 18 and 151, respectively. It is quit different the one which we defined in Fig. 4.

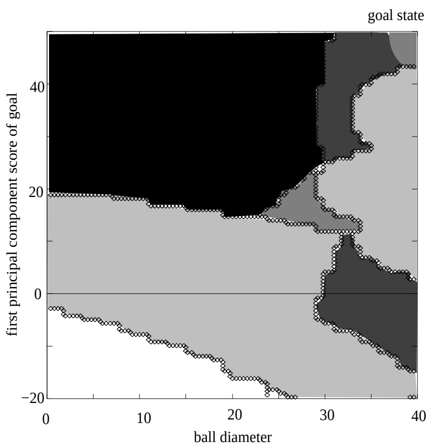


Figure 6: state space construction of real robot experiment

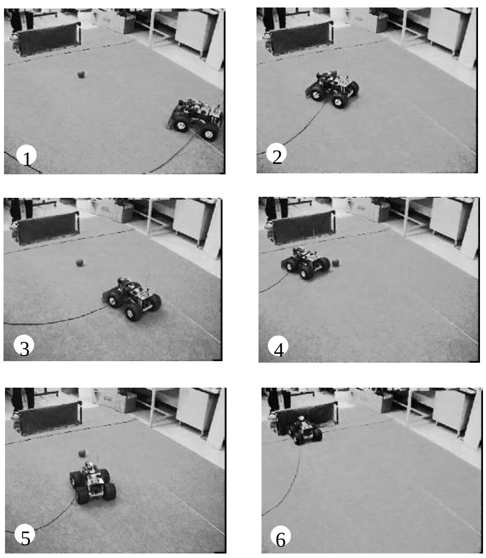


Figure 7: The robot succeeded in shooting a ball into the goal

5 Conclusions

The real robot has succeeded to acquire a behavior for a shooting a ball into a goal(Fig. 7). Acquiring various more complex task, such as pass a ball to a team mate or intercepting a ball from an opponent, is future work. Estimation of opponent models, and learning of team play are also future works.

In RoboCup-97, we will show the acquired behavior by learning.

References

[1] M. Asada, S. Noda, and K. Hosoda. Action-Based Sensor Space Categorization for Robot Learning. In *Proc. of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1502–1509, 1996.

[2] M. Asada, S. Noda, S. Tawaratsumida, and K. Hosoda. Vision-Based Reinforcement Learning for Purposive Behavior Acquisition. In *Proc. of IEEE International Conference on Robotics and Automation*, pp. 146–153, 1995.

[3] M. Asada, E. Uchibe, and K. Hosoda. Agents That Learn from Other Competitive Agents. In *Proc. of Machine Learning Conference Workshop on Agents That Learn from Other Agents*, 1995.

[4] M. Asada, E. Uchibe, S. Noda, S. Tawaratsumida, and K. Hosoda. Coordination Of Multiple Behaviors Acquired By A Vision-Based Reinforcement Learning. In *Proc. of the 1994 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2, pp. 917–924, 1994.

[5] J. H. Connel and S. Mahadevan. *Robot Learning.* Kluwer Academic Publishers, 1993.

[6] Y. Takahashi, M. Asada, and K. Hosoda. Reasonable Performance in Less Learning Time by Real Robot Based on Incremental State Space SegmentationAction-Based Sensor Space Categorization. In *Proc. of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1518–1524, 1996.

[7] E. Uchibe, M. Asada, and K. Hosoda. Behavior Coordination for a Mobile Robot Using Modular Reinforcement Learning. In *Proc. of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1329-1336, 1996.

[8] E. Uchibe, M. Asada, S. Noda, Y. Takahashi and K. Hosoda. Vision-Based Reinforcement Learning for RoboCup: Towards Real Robot Competition. In *Proc. of the IROS-96 Workshop on RoboCup*, pp.16–24, 1996.