International Journal of Fuzzy Logic and Intelligent Systems, vol. 6, no. 3, September 2006, pp. 264-269

Strategy of Object Search for Distributed Autonomous Robotic Systems

Ho-Duck Kim, Han-Ul Yoon, and Kwee-Bo Sim*

School of Electrical and Electronic Engineering, Chung-Ang University
221, Heukseok-Dong, Dongjak-Gu, Seoul, 156-756, Korea
Tel : +82-2-820-5319, Fax : +82-2-817-0553, E-mail : kbsim@cau.ac.kr

Abstract

This paper presents the strategy for searching a hidden object in an unknown area for using by multiple distributed autonomous robotic systems (DARS). To search the target in Markovian space, DARS should recognize their surrounding at where they are located and generate some rules to act upon by themselves. First of all, DARS obtain 6-distances from itself to environment by infrared sensor which are hexagonally allocated around itself. Second, it calculates 6-areas with those distances then take an action, i.e., turn and move toward where the widest space will be guaranteed. After the action is taken, the value of Q will be updated by relative formula at the state. We set up an experimental environment with five small mobile robots, obstacles, and a target object, and tried to research for a target object while navigating in a un known hallway where some obstacles were placed. In the end of this paper, we present the results of three algorithms – a random search, an area-based action making process to determine the next action of the robot and hexagon-based Q-learning to enhance the area-based action making process.

Key words : DARS, area-based action making, hexagon-based Q learning, object recognition

1. Introduction

Nowadays, robots are replacing human's work in dangerous fields, such as rescue jobs at fire-destroyed buildings or at gas-contaminated sites; information retrieval from deep seas or from space; and weather analysis at extremely cold areas like Antarctica. Multiple robots are especially needed to penetrate into hard-to-access areas, such as underground ant nests, to collect reliable and solid data. They can send data through cooperation and communication and can make independent decisions to act. Distributed autonomous robotic systems (DARS) are systems with multiple autonomous robotic agents, assigned with required functions. The most unique and important feature of DARS is that each system is a distributed system composed of multiple agents or robots [1]. With this feature, DARS can be applied for a wide range of application. DARS is now applied to multi-robot behavior, distributed control, coordinated control, cooperative operation, etc. DARS has received much attention since it can offer a new way of controlling multiple agents more flexibly and robustly. Parker proposed the heuristics approach algorithm for multiple robots and applied it to cleaning tasks [2]. Ogasawara used this system for several robots to transport a large object [3]. In this paper, we propose an area-based action making (AAM) process to control multiple robots against collision and guide individual robot to search through its own trajectory.

Reinforcement learning allows an agent to actively determine an action policy based on explorations of its environment. During exploration of an uncertain state space with reward, an agent can learn what to do by continuous tracking of its state history and appropriately propagating rewards through the state space [4]. In our research, we focused on Q-learning as a reinforcement learning technique. Because Q-learning is a simple way to solve Markovian action problems with incomplete information and on the basis of the action-value function Q that maps state-action pairs to expected returns [5]. In addition to this simplicity, Q-learning can adopt to the real world situation. For example, the state space can be matched with the physical space of the real world. An action also can be regarded as physical robot movement. In this paper, we propose the hexagon-based Q-learning to enhance the area-based action making process so that the learning process can better adapt to real world situations.

Typical mobile robot systems consist of a robot body (frame), vision system, sensor system, drive (motor) system, communication system, and main controllers. These systems can be applied in many ways, efficiently, according to the main task of the robots and to the parts of the robots to specially intellectualized [6]. Khepera is an example of a mobile robot. It consists of a main processor (Motorola 68331) driven by two

Manuscript received Aug. 30, 2006; revised Sep. 13, 2006.
* Corresponding author
This work was supported by the Korea Research Foundation Grant funded by the Korean Government(KRF-2005-042-D00268).

DC motors and eight infrared proximity sensors around its body (55mm in diameter and 30mm in height) [7]. It is still used in many fields, such as fuzzy control, wall following, and obstacle avoidance. The robot in our project looked similar to Khepera or RoboSot (one model of soccer robots).

In this paper, we propose an area-based action making (AAM) process and hexagon-based modified Q-learning to control multiple robots from collision and guide individual robot to search through its own trajectory. We organized DARS with five small mobile robots and sent them to a hallway, where some obstacles were lying about, to search for a hidden object. The organization of this paper is as follows. In chapter 2, our robot system and its functional block component are introduced. In chapter 3, the task of our DARS, the area-based action making algorithm, hexagon-based Q-learning adaptation, and object recognition with vision system are presented. In chapter 4, experimental results from the application of three different strategies to find the object are presented. In chapter 5, conclusions are presented.

2. Architecture of Small Mobile Robot

Our small mobile robot system which was made in lab consisted of four sub-parts and a main micro-controller part. The sub-parts were camera vision, sensor, motor, and Bluetooth communication module. Each sub-part had its own controller to perform its unique function more efficiently. The main micro-controller part controls the four sub-parts to avoid process collision and generates actions with the data from its sub-parts. Figure 1 shows the appearance, anatomy, and functional block diagram of the robot.

The main components of the robot are as follows. For the eye of the robot, Movicam II made by Kyosera is used. It is the CCD camera used for SKY cellular phone. Its size is 30×47×29 mm (width × length × height) and its weight is 12g, approximately. The robot has six infrared sensors, emitter and detector pairs, to measure the distance around itself. The emitter is Kodenshi EL-1kl3, high-power GaAs infrared sensor. The detector is ST-1kla, high sensitivity NPN silicon phototransistors. They are mounted in a durable and hermetically sealed TO-18 metal in a package. The six sensor pairs are placed at an angle of 60 degrees with one another to cover 360 degrees. NMB PG25L-024 stepping motor is used as the driving part. Its char acteristics are the following: drive voltage-12V, drive method 2-2 phase and 0.495° step angle. We used Bluetooth communication to make the robot very suitable for wireless communication systems.

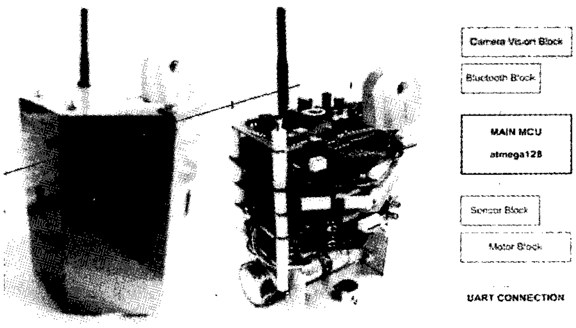

Fig. 1. Appearance (left), anatomy (center), and functional block diagram (right) of the robot.

The primary function of the main controller is to decide the subsequent actions with information, sent by the sensor and the camera vision block, and send the appropriate information to the motor block to control the direction of the robot. The control algorithm will be discussed in the next chapter.

3. Strategies of Object Search for DARS

The task of the robots is a follows: "Find the hidden object while tracking through an unknown hallway". To complete the task, the robots must be able to do the following. First, the robots should be able to search space as wide as they can. Second, they should avoid obstacles and not disturb each other during the task. Finally, they must recognize the object correctly To satisfy with these requirements, therefore, we need strategies to control the robots robustly and efficiently.

In this chapter, we introduce the three robot control strategies: area-based action making process, hexagon-based Q-Learning adaptation, and object recognition with vision system.

3.1. Area-based action making process

To search a more wide and safe space, we need to control the robots with a rule. We use the area-based action making (AAM) process, which is similar with the behavior-based direction change, to control the robots [8]. This process is referred to as AAM because each robot recognizes its surrounding not by 6-distances but its 6-areas. The key idea of the AAM process is to reduce the uncertainty of its surrounding. The robots recognize the shape of its surrounding, then take an action (turn and move forward) to where the widest space will be guaranteed. Consequently, each robot can avoid an obstacle and collision with other robots. Figure 2 depicts the different actions taken by distance-based action making (DAM) and by AAM in the same situation.

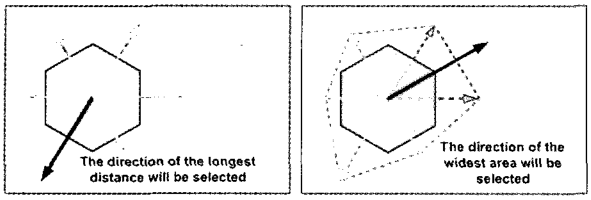

Fig. 2. The different action will be taken by DAM.

The advantage of AAM is illustrated by the following example. Figure 3 presents the result of each action making by DAM and AAM. In both case, the robot is surrounded by 4-obstacles. By DAM, the robot will be confused because it perceives that there is no obstacle in the southeast direction, and then it will try to keep tracking to the southeast. Finally, it will get stuck between two obstacles. This scenario is shown by the top picture in Fig. 3. By AAM, however, the robot will calculate the areas of its surrounding, and then it will recognize that an action to the northeast will guarantee the widest space. Therefore, the robot will change its direction to the northeast. This scenario is presented in the bottom picture in Fig. 3.

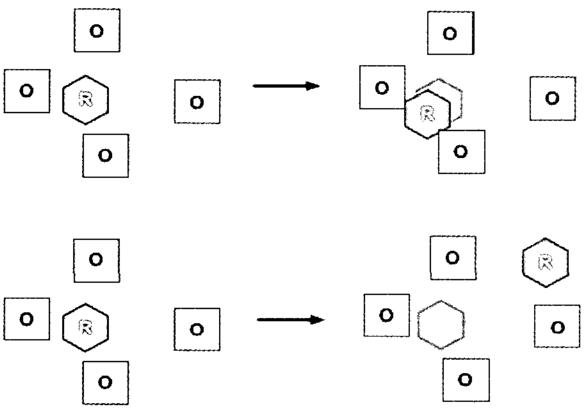

Fig. 3. An illustrative example of robot maneuvers by DAM(top) and by AAM(bottom).

In addition to the obstacle avoidance, AAM also make the robots to search their own space. This feature is advantageous when 2 or 3 robots meet at the same place. When they face each other, each robot will try to find more wide space. Consequently, the robot will change its direction to avoid the other robots and start to search in its own space again.

3.2. Hexagon-based Q-learning adaptation

Q-learning is a well-known algorithm for reinforcement learning. It leads the agent to acquire optimal control strategies from delayed rewards, even when the agent has no prior knowledge of the effects of its actions on the environment [9]. Q-learning algorithm is presented below.

For each s, a initialize the table entry Q(s,a) to zero
Observe the current state s
Do forever

- Select an action a and execute it
- Receive immediate reward r
- Observe the new state s'
- Update the table entry for $Q(s,a)$ as follows:

$$Q(s,a) \leftarrow r + \gamma \max_{a'} Q(s',a') \qquad (1)$$

- $s \leftarrow s'$

where s is a possible state, a is a possible action, r indicates the immediate reward value, and γ is the discount factor. Figure 4 is an illustrative example to explain Q-learning algorithm more clearly. Each grid square presents the possible states. The 'R' stands for a robot or agent. The values upon the arrows are relevant $\hat{Q}$ values with the state transition. For example, the value $\hat{Q}(s_1, a_{right}) = 72$, where a right refers to the action that moves R to its right [9]. If the robot takes the action to the right, the value will be updated for this entry

$$\begin{aligned} Q(s,a) &\leftarrow r + \gamma \max_{a'} Q(s',a') \\ &\leftarrow 0 + 0.9 \max\{63, 81, 100\} \\ &\leftarrow 90 \end{aligned} \qquad (2)$$

where $r = 0$, $\gamma = 0.9$ are predetermined values.

The Q-learning for our robot system was adapted to enhance the AAM process. The adaptation can be performed with a simple and easy modification, named hexagon-based Q-learning Figure 5 is an illustrative example of hexagon-based Q-learning. In Fig. 5, intuitively, we know that the only thing that was changed is the shape of state space. We changed the shape of the space, from a square to a hexagon, so that the robot can recognize its surrounding by 6-areas. According to this adaptation, the robot takes an action to 6-direction and has 6-table entry $\hat{Q}$ value. In the left of Fig. 5, the robot is in the initial state. Now, if the robot decides that +60 degree guarantee the widest space after calculation of its 6-areas of surrounding, the action of the robot would be a_{+60}. After the action is taken, if $Area3$ is the widest area, the value of $\hat{Q}(s_1, a_{+60})$ will be updated by the formula (1) in the Q-learning algorithm as follows

$$\begin{aligned} Q(s_1, a_{init}) &\leftarrow r + \gamma \max_{a'} Q(s_2, a_{+60}) \\ &\leftarrow 0 + \gamma \max\{Area1, Area2, \ldots, Area6\} \\ &\leftarrow \gamma Area3 \end{aligned} \qquad (3)$$

where 0 is the predetermined immediate reward. After the movement from the initial state to the 1st next state, immediate reward becomes the difference between *the sum of total area before action is taken* and *the sum of total area after action is taken*. Thus,

$$r = \sum_{j=1}^{6} Area_j - \sum_{i=1}^{6} Area_i \tag{4}$$

where $Area_i \in s_1$ and $Area_j \in s_2$, respectively.

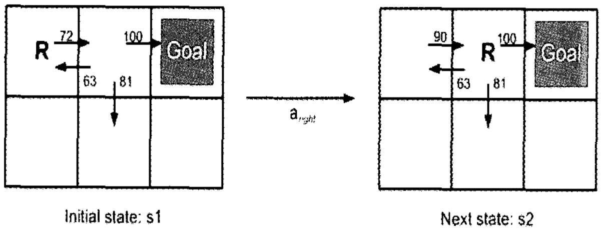

Fig. 4. Illustrative example about Q-learning.

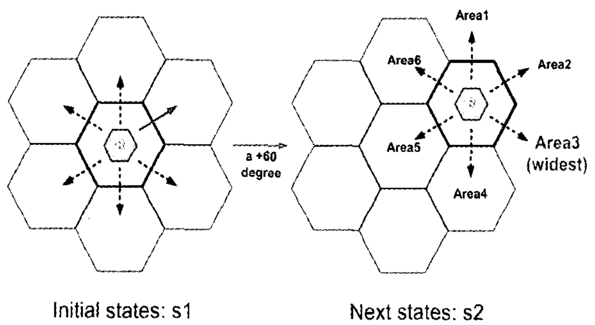

Fig. 5. Hexagon-based modified Q-learning.

Ultimately, the robot can determine its trajectory by learning this $\hat{Q}$ value. In the real world experiment, however, battery consumption is a problem. If the robot has to perform infinite iterations to complete task, total system will fail. Therefore, a system must be set up to cancel the former action and move back to the earlier state, when the former action causes any bad reward or result.

3.3. Object recognition with vision system

The 5-robots, which try to search the object, should recognize the object correctly by the object's color and shape. We set up the color of the object as green and that of 5-robots as orange. The object was a stationary robot having the same shape. It was a located at a hidden place near the obstacle.

The 5-robots will have to decide whether they have finished the task by detecting the object after each action is taken. The image processing should be quick and robust so that no portion of receiving image data is missed and the situation is not judged wrongly. However, the total image size from the camera mounted on the robot is rather huge to process within a short time. Therefore, the image size must be reduced by extraction the area of interest to speed up the image processing. The size of interested area extracted by program was 160×120 bytes from the center of image. It is presented in Fig. 6. When the robot is close enough to the object, the interested area will be filled with green. If the sum of the green values is greater than a threshold value predetermined in the program, then the robot accepts this situation as object detection.

When the robot detects the object, it turns on a green LED, which is attached on the top of the robot. Otherwise, it turns on a red LED. After turning on the green LED, the robot starts to approach to the object. Finally, if the robot is blinking both LEDs alternately, the task is terminated.

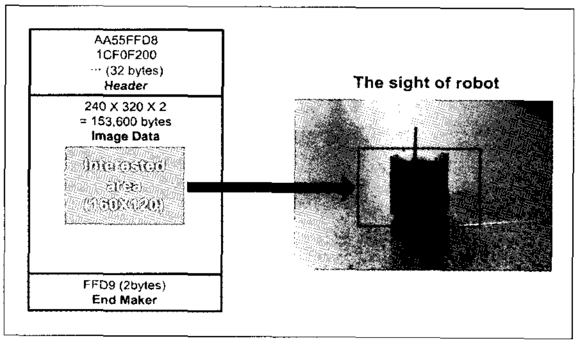

Fig. 6. Area if interest extracted from the image data.

4. Experiments with Three Different Control Algorithms

Experiments were performed by using three different control methods: random search, AAM, and enhanced AAM by hexagon-based Q-learning.

First, we used the random search control method to find the hidden object. The main controller generated a random number and decided the next action corresponding to this number. Random search is not so strong method to control the robot efficiently. Therefore, random did not perform well. Moreover, it is very time and power consuming in the real world situation. The result showed that random search is a horrible method to adopt to a real robot system. In Fig. 7, the white arrow points out the object (same in Fig. 8 and Fig. 9). During random search, even though the robots are within a close distance to the search object, some robots failed to find the object.

Second, we applied AAM to the robots. With the feature of AAM, the robots sense their environment by 6-infrared sensors and calculate 6-area with these values. When the calculation is done, each robot tries to move to where the widest area will be guaranteed. In our 2nd experiment, after the robots started to move, each robot spread out into the environment. Consequently, the AAM performed better than random search. Figure 8 shows that the two robots, which is located the right side of the object, succeeded to complete the task. These two robots are designated by black arrow in Fig. 8.

Finally, we adopted the hexagon-based Q-learning to AAM as a modified control method. This method allowed the robots to reduce the probability of wrong judgment and compensated wrong judgment by reinforcement learning. Each robot tried to search its own area as in the 2nd experiment, however, it canceled the decided action if the action caused negative (or bad) immediate reward value. By using the hexagon-based Q-learning adaptation to AAM, more than 2 robots completed the task during the 10-trials.

Fig. 7. Five-robots are searching the object using random search.

Fig. 8. Search by area-based action making.

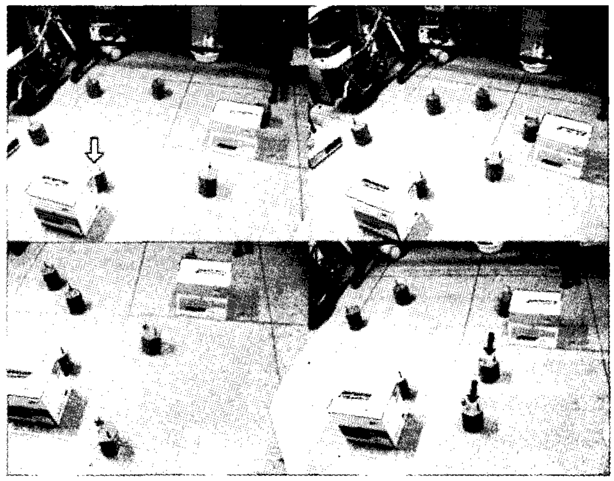

Fig. 9. Search with adaptation hexagon-based Q-learning to AAM

The results of our experiment are presented in Fig. 10. With random search, one robot found the object at the 2nd trial and 6th trial, although these detections can be considered as just coincidence. Therefore, we can say the random search has no remarkable meaning. With AAM, the robots performed better than with random search, with the average performance above 1 during the all trial. Finally, with the adaptation of hexagon-based Q-learning to AAM, the results were remarkable. Especially, 3-robots succeeded to find the object at the 4th, 6th, 8th, 9th, and 10th trial.

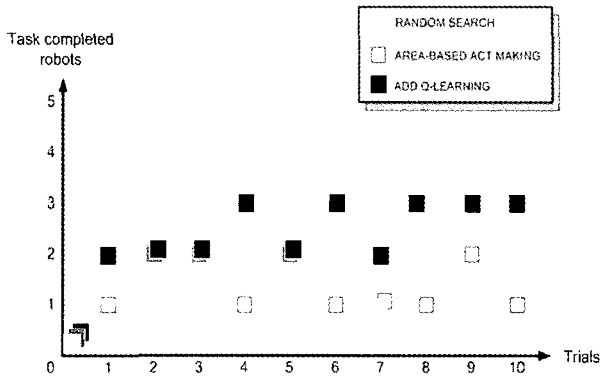

Fig. 10. Experimental result with three different control algorithms.

5. Conclusions

In this paper, we proposed the area-based action making process and hexagon-based Q-learning modification to search objects, hidden in unknown space, for 5 of our self-made small mobile robots. The experimental results from the application of the three different control methods in the same environmental situations were presented. The area-based action making process and hexagon-based Q-learning modification can be a new way for robot object search in unknown space. These two algorithms also make the agents avoid obstacles during their search. In our research, first, we need to clarify the problem of accessing to the object. This means that if multiple robots are to carry out a task such as object transporting or block stacking, the robots need to recognize the object then approach to it. Therefore, we need to develop the robust accessing algorithm. Naturally, some grippers need to be attached to both sides of the robot. Second, our robot systems should be improved so that the main part and the sub-parts adhere more strongly. In addition, stronger complex algorithms such as Bayesian learning or TD(λ) method should be adapted. Third, a self-organizing Bluetooth communication network should be built so that robots can communicate with each other robustly even if one or more robots are lost. Finally, the total system should be refined.

References

[1] H. Asama, T. Fukuda, T. Arai, and I. Endo, *Distributed autonomous robotic systems*, Springer-Verlag, 1994.

[2] L. Parker, "Adaptive action selection for cooperative agent teams," *Proc. 2nd. Inter- national Conf. on Simulation of Adaptive Behavior*, pp.442-450, 1992.

[3] G. Ogasawara, T. Omata, T. Sato, "Multiple movers using distributed, decision-theoretic control," *Proc. of Japan-USA Symposium On Flexible Automation*, Vol. 1,pp. 623-630, 1992.

[4] D. Ballard, *An introduction to natural computation*, The MIT Press, 1997.

[5] J. Jang, C. Sun, E. Mizutani, *Neuro-Fuzzy and soft computing*, Prentice-Hall, 1997

[6] J. H. Kim, H. S. Sim, S. H. Kim, *Robot soccer engineering*, Dooyang, 2003.

[7] S. H. Lian, "Fuzzy logic control of an obstacle avoidance robot," *Proc. of the Fifth IEEE International Conf. on Fuzzy Systems*, vol. 1, pp.26-30, 1996.

[8] W. Ashley, S. Balch, and T. Balch, "Value-based observation with robot teams (VBORT) using probabilistic techniques," *Proc. of ICAR 2003*, vol. 1, pp. 1-8, 2003.

[9] T. Mitchell, *Machine Learning*, McGraw-Hill, 1997.

Ho-Duck Kim received the B.S. degrees in the School of Electrical and Electronics Engineering from Chung-Ang University, Korea, in 2005. He is currently pursuing the M.S. degree at the Chung-Ang University. His research interests include Evolvable H/W Processor Design, Embedded Linux System, Genetic Algorithm, DARS, etc.

Han-Ul Yoon received the B.S. and M.S. degrees in the School of Electrical and Electronics Engineering from Chung-Ang University, Korea in 2004 and 2006 respectively. He is currently pursuing the Ph. D. degree at the University of Illinois at Urbana-Champaign. His research interests include Evolvable H/W, Processor Design, Genetic Algorithm, DARS, etc.

Kwee-Bo Sim

Kwee-Bo Sim received his B.S. and M.S. degrees in the Department of Electronic Engineering from Chung-Ang University, Korea, in 1984 and 1986 respectively, and Ph.D. degree in the Department of Electrical Engineering from The University of Tokyo, Japan, in 1990. Since 1991, he has been a faculty member of the School of Electrical and Electronics Engineering at Chung-Ang University, where he is currently a Professor. His research interests are in artificial life, emotion recognition, ubiquitous intelligent robot, intelligent System, computational intelligence, intelligent home and home network, ubiquitous computing and Sense Network, adaptation and machine learning algorithms, neural network, fuzzy system, evolutionary computation, multi-agent and distributed autonomous robotic system, artificial immune system, evolvable hardware and embedded system etc. He is a member of IEEE, SICE, RSJ, KITE, KIEE, KFIS, and ICASE Fellow. He is currently President of the KFIS.

Phone : +82-2-820-5319

Fax : +82-2-817-0553

E-mail : kbsim@cau.ac.kr