

다수 로봇의 목표물 탐색을 위한 Area-Based Q-learning 알고리즘

Area-Based Q-learning Algorithm to Search Target Object of Multiple Robots

윤한얼 · 심귀보

Han-UI Yoon and Kwee-Bo Sim

중앙대학교 전자전기공학부

요 약

본 논문에서는 다수 로봇의 목표물 탐색을 위한 area-based Q-learning 알고리즘에 대해 논한다. 선형적 정보가 없는 공간 내의 목표물을 탐색하기 위해, 로봇들은 주위의 환경을 인식하고 스스로 다음 행동에 대한 결정을 내릴 수 있어야 한다. Area-based Q-learning에서, 먼저 각 로봇은 정육각형을 이루도록 배치된 6개의 적외선 센서를 통해 자신과 주변 환경 사이의 거리를 구한다. 다음으로 이 거리데이터들로부터 6방향의 면적(area)을 계산하여, 보다 넓은 행동반경을 보장해주는 영역으로 이동(action)한다. 이동 후 다시 6방향의 면적을 계산, 이전 상태에서의 이동에 대한 Q-value를 업데이트 한다. 본 논문의 실험에서는 5대의 로봇을 이용하여 선형적 지식이 없고, 장애물이 놓여 있는 공간에서의 목표물 탐색을 시도하였다. 결론에서는 3개의 제어 알고리즘 - 랜덤 탐색, area-based action making (ABAM), hexagonal area-based Q-learning - 을 이용하여 목표물 탐색을 시도한 결과를 보인다.

Abstract

In this paper, we present the area-based Q-learning to search a target object using multiple robot. To search the target in Markovian space, the robots should recognize their surrounding at where they are located and generate some rules to act upon by themselves. Under area-based Q-learning, a robot, first of all, obtains 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 search for a target object while navigating in a unknown hallway where some obstacles were placed. In the end of this paper, we presents the results of three algorithms - a random search, area-based action making (ABAM), and hexagonal area-based Q-learning.

Key words : area-based action making, Q-learning, hexagonal area-based Q-learning

1. 서 론

최근 들어 화재가 발생한 건물에서의 구조 활동, 가스 누출 사고 지역의 정보 수집, 깊은 바다 속의 탐색, 극지방과 같은 곳에서의 기후 조사와 같은 부분에서, 로봇이 사람을 대신하여 작업을 수행하고 있다. 특별히, 땅 아래 곤충의 집과 같은 사람이 직접 접근하기 힘든 곳의 탐색에서, 신뢰성과 이용가치가 높은 정보의 획득을 위해 다수의 소형 로봇들이 보내진다. 다수의 로봇을 보다 유연하고 강인하게 제어하기 위한 방법은 현재까지 많은 주목을 받아왔다. Parker는 다수 로봇의 작업 수행을 위해 heuristic 형태의 알고리즘을 제안하였다 [1]. Ogasawara는 다수의 로봇을 이용해 커다란

물체를 수송하기 위해 자율 분산 로봇 제어 방식을 이용하였다 [2]. 본 논문에서는 다수의 로봇이 어떤 작업을 수행함에 있어 서로간의 충돌을 피하고, 자신만의 고유한 영역을 탐색하도록 하기 위한 방법으로 area-based action making을 제안한다. 이 area-based action making은 hexagonal area-based Q-Learning의 바탕이 된다.

강화 학습은 agent로 하여금 주변 환경의 탐색을 통해 능동적으로 환경에 대한 행동을 결정하도록 한다. 보상값이 존재하는 어떤 불확실한 영역을 탐색하는 동안 agent는 연속적인 상태 공간을 따라 적절한 보상값을 전달함으로써, 임의의 상태에 대해 어떠한 행동을 취해야 할지를 학습하게 된다 [3]. 강화 학습을 구현하기 위한 많은 방법 중, 본 논문에서는 Q-learning을 이용하였다. 그 이유는 Q-learning은 불완전한 정보를 가진 Markovian 공간에서의 행동 결정에 대해, 어떤 상태와 행동으로 이루어진 Q-함수를 기본으로 하여 문제의 해결에 쉬운 방법을 제공하기 때문이다 [4]. 또한 이 임의의 상태 공간을 실제로 물리적인 공간으로 간주될 수 있다. 본 논문에서는 area-based action making을 강화하기 위해 hexagonal area-based Q-learning을 적용한다.

접수일자 : 2005년 4월 1일

완료일자 : 2005년 5월 16일

감사의 글 : 이 논문은 2005년도 중앙대학교 학술연구비(일반연구비) 지원에 의한 것입니다. 연구비 지원에 감사드립니다.

본 논문의 2장에서는 area-based action making에 대해 나타낸다. 3장에서는 hexagonal area-based Q-learning에 대해 논한다. 4장에서는 로봇들이 위치를 파악하기 위한 좌표 시스템에 대해 설명한다. 5장에서는 위의 3가지 제어 알고리즘들을 적용한 목표물 탐색의 실험 결과를 보인다. 마지막으로 6장에서는 결론 및 향후 과제에 대해 논한다.

2. Area-Based Action Making Process

Area-based action making(ABAM)은 로봇이 어떤 상태에서 환경에 대한 정보를 획득한 후, 다음 행동을 결정하는데 기본 바탕이 되는 방법이다. 이를 ABAM이라 부르는 이유는, 로봇이 자신 주변의 환경을 둘 사이의 거리가 아닌 자신 주변의 면적을 계산하여 얻어진 정보로부터 다음 행동을 결정하기 때문이다. 이 방식의 핵심은 로봇으로 하여금 자신 주위의 불확실성을 줄여 나가도록 한다는 데 있다. 따라서 ABAM은 행동 기반 방향 전환 (behavior-based direction change) 방식과 많은 유사점을 가지고 있다고 할 수 있다 [5][6]. ABAM에서 로봇은 자신 주변 공간의 형태를 파악하고, 보다 넓은 영역을 보장해주는 곳으로 자신의 위치를 이동시킨다. 그림 1은 센서를 통해 획득한 거리 정보를 이용 (distance-based action making: DBAM)할 때와 ABAM을 이용할 때 로봇의 행동 결정의 차이점을 보여 준다. 본 논문에서는 실험을 위해 자체 제작된 소형 로봇을 이용하였고, 360°를 모두 탐지 할 수 있도록 6개의 적외선 센서를 60°간격으로 로봇 주위에 배치하였다.

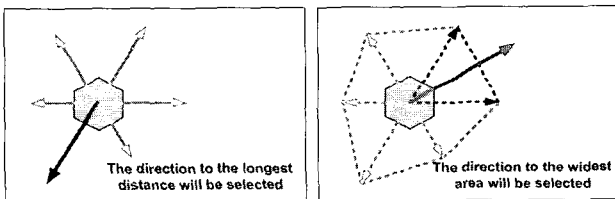


그림 1. 거리 기반 방식과 면적 기반 방식에 의한 로봇의 행동 결정.
Fig. 1. Different action taken under DBAM and ABAM.

그림 2는 같은 환경에 있는 로봇들의 서로 다른 다음 행동 결정의 예를 통해, ABAM의 장점을 보여준다. 그림에서 로봇은 4대의 장애물에 둘러싸여 있다. 왼쪽 그림과 같이 DBAM을 따를 경우, 로봇은 남서쪽(+240°)에 장애물이 없다고 판단하여 계속해서 그 방향으로 진로를 결정할 것이다. 결국 로봇은 두 장애물 사이에서 벗어나기 어렵게 된다. 반면에 오른쪽과 같이 ABAM을 이용할 경우, 로봇은 보다 넓은 영역을 보장해 주는 방향으로 행동을 취하므로 그림의 오른쪽과 같이 로봇은 장애물에 둘러싸인 환경을 빠져나올 수 있다.

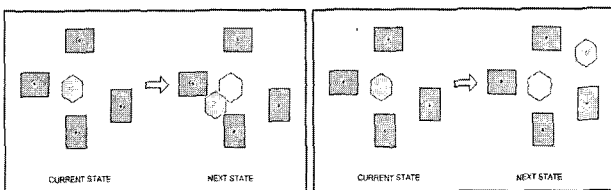


그림 2. DBAM과 ABAM에 따른 행동의 예.
Fig. 2. Illustrative examples of robot maneuvers.

ABAM을 이용한 목표물 탐색 과정을 쉽게 설명하기 위해 다음과 같은 상황을 가정해 보자. 다수개의 컴퓨터를 이용하여 방대한 크기의 메모리로부터 (예: DB 서버) 원하는 데이터를 검색하려 한다. 모든 데이터들은 규칙적으로 저장되지 않았으며, 정렬되지도 않았다. 또한 특수한 작업을 위해 예약되어 있는 메모리 공간도 있다고 가정하자. 이 경우, 가장 빠르게 원하는 데이터를 찾는 방법은 각각의 컴퓨터에 일정한 크기의 영역을 할당 한 후, 그 할당 된 영역을 검색하게 하는 것이다. 흥미롭게도, 다수의 컴퓨터에 각각의 탐색 영역을 할당 하는 것은 다수의 로봇을 정해진 탐색 공간에 ABAM을 이용하여 풀어놓는 것과 같다. 데이터 탐색을 위한 포인터의 움직임은 로봇의 이동경로라 생각할 수 있다. 또, 예약되어 있는 메모리 공간은 로봇의 탐색 공간상에서 장애물이 놓여 있는 공간이라 볼 수 있다. 따라서 ABAM을 적용한 다수개의 로봇을 통해 어떠한 공간에서 목표물을 탐색하는 것은 멀티프로세싱을 이용하여 전체 프로세스의 성능 시간을 줄이는 것과 유사하다 [7]. 그림 3에 위 두 경우에 대한 개념적 유사성을 나타내었다.

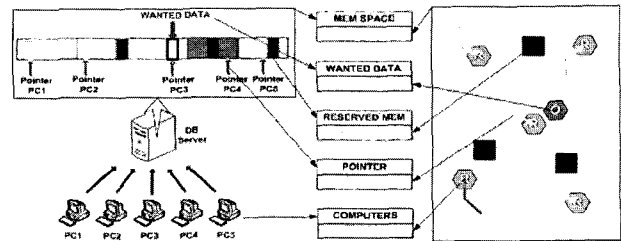


그림 3. Multiprocessing을 통한 데이터 탐색과 다수의 로봇을 이용한 목표물 탐색의 유사성.

Fig. 3. Similarity between data search using multiple computers and an object search using multiple robots.

3. Hexagonal Area-Based Q-Learning

Q-learning은 대표적인 강화 학습 알고리즘 중의 하나이다. Q-learning은 agent가 환경에 대한 선행적 정보를 가지고 있지 않을 때에도, 행동에 대한 보상값을 통해 최적의 행동 전략을 획득할 수 있도록 해준다 [8][9]. Q-learning 알고리즘을 표 1에 나타내었다. 여기서 s 는 상태를, a 는 행동을, r 은 보상값을, γ 는 Q-함수값의 조정을 위한 계수(discount factor)이다.

표 1. Q-learning 알고리즘
Table 1. Q-learning algorithm

For each s, a initialize the table entry $Q(s, a)$ zero	
Observe the current state s	
Do forever	
• Select the action a and execute it	
• Receive the immediate reward r	
• Update the table entry for $Q(s, a)$	(1)
• $s \leftarrow s'$	

그림 4는 Q-learning에 대한 실제적인 예를 보여준다. 각

각의 정사각형은 상태를 나타낸다. R은 로봇을 나타낸다. 상태의 천이에 따른 화살표 위에 나타난 값은 그 행동을 취함에 따른 Q값을 나타낸다. 예를 들어, 초기 상태에서 오른쪽으로 상태를 천이하는데 따른 Q값은 화살표 위의 값인 $Q(s_1, a_{right}) = 72$ 와 같다.

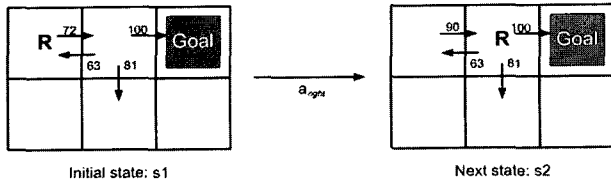


그림 4. Q-learning의 예.
Fig. 4. An illustrative example of Q-learning.

초기 상태에서 만약 로봇이 오른쪽으로 행동을 취한다면, 업데이트 되는 Q값은 $r=0, \gamma=0$ 을 초기값으로 할 때

$$\begin{aligned}
 Q(s_1, a_{+60^\circ}) &\leftarrow r + \gamma \max_{a'} Q(s', a') \\
 &\leftarrow 0 + 0.9 \max_{a'} \{63, 81, 100\} \\
 &\leftarrow 90
 \end{aligned} \quad (2)$$

이 된다.

본 논문에서는 면적 기반 행동 결정을 강화하기 위해 Q-learning을 사용하였다. 기반이 되는 단위 면적은 6각형으로 하였다. 그 이유로는, 단위 면적에 조합에 따라 무한히 확장될 수 있는 다각형은 4각형을 제외하면 6각형뿐이기 때문이다. 또한 6각형을 단위 면적으로 하면 더 정밀한 로봇 주위 환경의 정보 수집이 가능하다. 따라서 로봇은 6개의 적외선 발/수 센서를 주위에 가지고 있다. 그림 5에 hexagonal area-based Q-learning 적용의 간단한 예를 나타내었다. 만약 초기 상태에서 로봇이 $+60^\circ$ 방향으로 행동을 결정하였고, 행동을 실행한 후 Area6'가 가장 넓은 영역을 보장하였다고 하면, 이에 대한 Q값은 $Q(s_1, a_{+60^\circ})$ 로 표현될 수 있고, 그 값은 초기 상태에서의 보상값이 0 이라면,

$$\begin{aligned}
 Q(s_1, a_{+60^\circ}) &\leftarrow r + \gamma \max_{a'} Q(s', a_\theta') \\
 &\leftarrow 0 + \gamma \max_{a'} \{Area1', Area2', \dots, Area6'\} \\
 &\leftarrow \gamma Area6'
 \end{aligned} \quad (3)$$

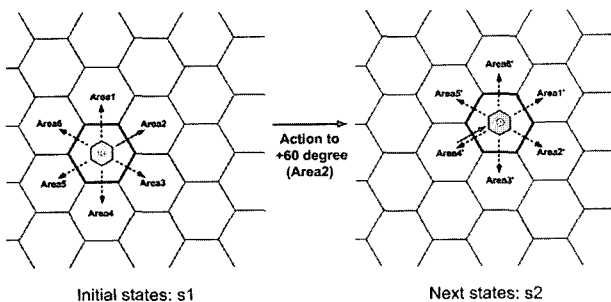


그림 5. Hexagonal area-based Q-learning.
Fig. 5. Hexagonal area-based Q-learning.

와 같이 된다.

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

단, $Area_i \in s$ 이고 $Area_j \in s'$ 이다.

결과적으로 로봇은 이 Q값을 학습함으로써 진행 경로를 결정할 수 있다. 그러나 소프트웨어 상에서의 무한 반복은 현실 세계에서는 배터리 소비 문제를 고려할 때 구현이 불가능 하다 할 수 있다. 따라서 본 논문의 실험에서는 이전 상태에서 결정된 행동이 심각한 결과를 초래하였다면 다시 이전의 상태로 돌아가도록 해주었다. 전체 과정을 표 2에 나타내었다.

표 2. Hexagonal area-based Q-learning 알고리즘.

Table 2. Hexagonal area-based Q-learning algorithm.

For each s, a initialize the table entry $Q(s, a)$ zero
Calculate 6-areas at the current state s
Do until the task is completed
• Select the action a_θ to the widest area, and execute it
• Receive the immediate reward r
• Observe the new state s'
If $Q(s', a_\theta')$ is greater or equal than $Q(s, a_\theta)$
• Update the table entry for $Q(s, a_\theta)$
• $s \leftarrow s'$
Else, if $Q(s', a_\theta')$ is far less than $Q(s, a_\theta)$
• Move back to the previous state
• $s \leftarrow s'$

4. Hexagonal area-based Q-learning에서 로봇의 위치 표현을 위한 좌표 시스템

목표물을 발견했을 때의 위치를 저장하고, 또 상호 협조 작업의 수행을 위해, 로봇의 위치를 나타낼 수 있는 좌표 시스템이 요구된다. 앞 절에서 설명한 hexagonal area-based Q-learning의 특성 상, 6개의 면적은 6개의 정삼각형으로 나눌 수 있으므로 X-Y 좌표계를 바탕으로 매우 유용한 좌표 시스템을 만들 수 있다.

그림 6은 초기 상태에서 로봇이 결정할 수 있는 행동(action)중 하나를 보여준다. 두 육각형의 중심을 연결하는 가장 긴 변의 길이를 $2R$ 이라 하면 삼각함수의 공식에 따라, 밑변과 높이는 각각 $\sqrt{3}R$ 과 R 로 나타낼 수 있다. 따라서 기하학적인 해석을 통해, 로봇은 $+0^\circ, +60^\circ, -60^\circ$ 의 행동에 따라 각각 $0, \sqrt{3}R, -\sqrt{3}R$ 의 x 변위를 갖는다. 또 y 변위에 대해서는 $+0^\circ, \pm 60^\circ, \pm 120^\circ$ 의 행동에 대해 각각 $+2R, +R, -R$ 의 값을 갖는다. $\pm 180^\circ$ 로의 행동은 앞 절에서 설명한 심각한 Q값의 저하에 따른, 상태 및 행동을 취소하는 경우에 해당하므로 좌표값의 업데이트가 요구되지 않는다. 이와 같이 hexagonal area-based Q-learning의 특성에 의해 간단하고도 튼튼한 좌표 시스템을 생성할 수 있다. 표 3에 N 번째 상태에서의 로봇의 행동 $\theta_{s[N]}$ 에 따른 x, y 변위를 정리해 놓았다.

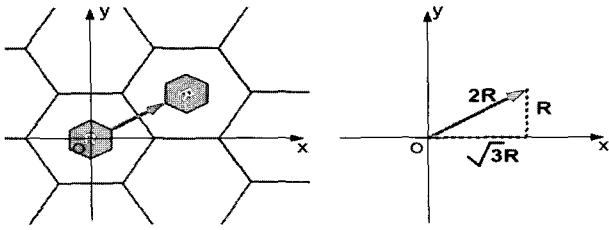


그림 6. 로봇의 이동에 대한 기하학적 해석.

Fig. 6. Geometrical vector analysis for robot movement.

표 3. 로봇의 행동 $\theta_{S[M]}$ 에 따른 x, y 변위.

Table 3. The displacement of x and y under robot's action.

$\theta_{S[M]}$	x 변위	y 변위
0	0	$2R$
+60	$\sqrt{3}R$	R
-60	$-\sqrt{3}R$	R
+120	$\sqrt{3}R$	$-R$
-120	$-\sqrt{3}R$	$-R$
± 180	0	0

그림 7은 로봇이 3번의 행동을 취할 동안 현재의 위치를 계산하는 과정을 나타낸다. 로봇의 현재 위치는 취한 행동에 따른 x, y 변위를, 벡터 덧셈에 의해 구한다. 정북쪽을 기준으로, 오른쪽을 각 θ 의 +변위로, 왼쪽을 -변위로 잡는다. 초기상태 $S[0]$ 에서 $\theta_{S[0]} = 0^\circ$ 이다. 이것은 현재 로봇이 정북쪽을 향하고 있음을 의미한다. 그림에서, 로봇이 상태 $S[0]$ 에서 $S[1]$ 으로 행동을 취한다면 $\theta_{S[1]} = +60^\circ$ 가 되고, 이것은 로봇의 현재 진행방향이 $+60^\circ$ 임을 의미한다. 다음으로 로봇이 상태 $S[1]$ 에서 $S[2]$ 로 행동을 취한다면 $\theta_{S[2]} = +60^\circ + (-60^\circ) = 0^\circ$ 와 같이 된다. 따라서 $\theta_{S[M]}$ 을 일반식으로 나타내면 다음과 같다.

$$\theta_{S[M]} = \sum_{k=0}^N \alpha_k \quad (5)$$

$$= \begin{cases} \alpha_0 = 0, \\ \alpha_1, \\ \alpha_k + \theta_{S[k-1]} & \text{if } 1 < k \leq N \end{cases}$$

여기서 α_k 는 그림 7에서 Y축과 로봇의 현재 진행 방향 사이의 각을 나타낸다. 결론적으로 로봇의 현재 위치는 다음과 같은 과정에 의해 구해진다.

- ABAM으로부터 행동($\theta_{S[M]}$)을 결정한다.
- $\theta_{S[M]}$ 에 따른 x, y 변위를 Look-up table로부터 읽어온다.
- $x \leftarrow x + x$ 변위
- $y \leftarrow y + y$ 변위

예를 들어, 그림 7에서 $\theta_{S[1]} = +60^\circ$ 이므로 로봇의 현재위치 $P_{S[1]}$ 는 $(x, y) = (\sqrt{3}R, R)$ 이다. 로봇이 $S[2]$ 로 행동을 취한다면 식 5에 따라, $\theta_{S[2]} = \alpha_1 + (-60^\circ) = +60^\circ + (-60^\circ) = 0^\circ$ 이고 $P_{S[2]}$ 은 $(x, y) = (\sqrt{3}R + 0, R + 2R) = (\sqrt{3}R, 3R)$ 과 같이 된다. 마지막으로 $\theta_{S[3]}$ 는 $\theta_{S[3]} = \alpha_2 + (-60^\circ) =$

$0^\circ + (+60^\circ) = 60^\circ$ 가 되고 $P_{S[3]}$ 는 $(x, y) = (\sqrt{3}R + \sqrt{3}R, 3R + 3R) = (2\sqrt{3}R, 2R)$ 이 된다. 최종적으로 목표물을 탐색하는데 성공한 로봇의 현재 좌표는 목표물과 바로 인접한 육각형 중 하나의 위치가 된다.

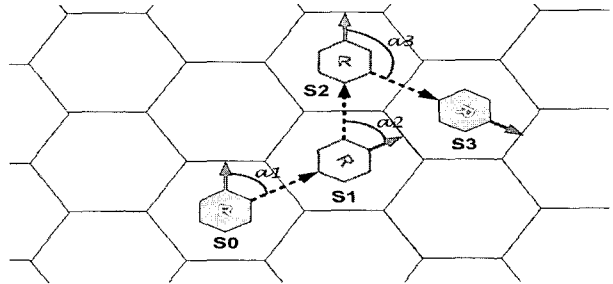


그림 7. 로봇의 위치 계산에 대한 예.

Fig. 7. An illustrative example of calculating location.

5. 다수 로봇을 이용한 목표물 탐색 실험

본 논문의 실험에서는 5대의 로봇에 서로 다른 3종류의 탐색 알고리즘-랜덤 탐색, ABAM, Hexagonal area-based Q-learning-을 적용하여 선형적 지식이 없는 환경에서 장애물 뒤에 숨어있는 목표물 탐색을 시도하였다. 5.1절에서는 실험에 사용된 로봇의 구성에 대해 보이고 5.2절에서는 3종류의 알고리즘을 각각 적용했을 때의 결과를 보인다.

5.1. 소형 이동 로봇의 구성

실험에 이용된 소형 이동 로봇은 1개의 메인부 마이크로 컨트롤러가 UART 커넥션으로 연결된 4개의 서브모듈 컨트롤러를 제어하는 방식으로 되어있다. 4개의 서브모듈은 각각 카메라-비전부, 센서부, 모터 구동부, 블루투스 통신부 로 구성된다. 메인부 컨트롤러는 서브모듈 컨트롤러로부터 받은 데이터를 바탕으로 로봇의 움직임을 제어한다. 그림 8은 로봇의 외형, 내부 구성, 기능별 블록 다이어그램을 보여준다.

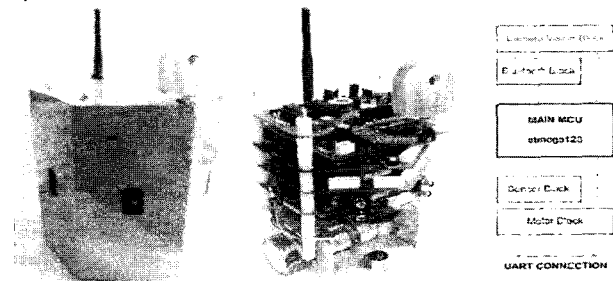


그림 8. 실험에 이용된 소형 이동 로봇.

Fig. 8. Small mobile robot employed in the experiments.

로봇의 눈으로 사용된 카메라는 Kyosera사의 Movcam II 이다. 휴대폰에 사용되는 CCD 카메라로 $30 \times 16.4 \times 47$ (가로×세로×높이) mm의 사이즈를 갖는다. 센서부는 주위 360도를 모두 커버하기 위해 60° 간격으로 배치된 6개의 발/수 적외선 센서로 구성되어 있다. 발광부는 Kodenshi사의 EL-1kl3 GaAs 적외선 센서를 사용하였고, 수광부는 ST-1kla NPN silicon 포토-트랜지스터를 사용하였다. 로봇

의 발인 구동부는 NMB PG25L-024 스테핑 모터 2조로 이루어져 있다. 이 모터는 2상 2여자 방식으로 구동하고 0.495°의 스텝각을 갖는다. 로봇의 꼭대기에는 PC와의 통신을 위해 블루투스 모듈을 장착하였다 [10][11].

5.2. 목표물 탐색 실험

5대의 로봇들이 수행할 작업은 “선행적 지식이 없는 환경에서 장애물 사이에 숨어있는 목표물을 탐색” 하는 것이다. 실험에서 탐색하는 로봇들은 오렌지색, 목표물은 녹색으로 정해주었다. 목표물은 정지해있는 로봇이며 색깔을 제외하곤 다른 로봇들과 외형상 동일한 구조를 가진다. 로봇들은 목표물의 색깔과 모양을 인식하여 작업을 계속 수행할 것인지 작업이 완료할 것인지를 결정하게 된다.

5.2.1 랜덤 탐색

첫 번째로 랜덤 탐색 알고리즘을 적용하여 목표물 탐색을 시도하였다. 메인 컨트롤러는 서브모듈로 부터의 데이터 수집없이 자체적으로 난수를 발생하여 그 수에 해당하는 행동을 취하게 된다. 랜덤 탐색은 평균 이하의 성능을 나타냈을 뿐만 아니라 배터리의 소모도 매우 크게 나타났다. 또한 대부분의 경우 5대의 로봇 모두가 목표물 탐색에 실패하였다. 따라서 랜덤 탐색은 실제 소형 로봇에 적용하기에 적당하지 못한 알고리즘임을 알 수 있다. 그림 9는 5대의 로봇이 랜덤 탐색을 이용하여 목표물 (흰색 화살표로 표시된 것)을 탐색하고 있는 과정을 보여준다.

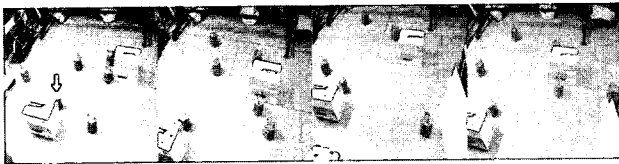


그림 9. 랜덤 탐색을 이용한 목표물 탐색.

Fig. 9. Five-robots are searching the object using random search.

5.2.2 ABAM을 이용한 탐색

두 번째로 ABAM을 적용하여 목표물 탐색을 시도하였다. ABAM에 의해 로봇은 6면적을 계산하여 자신 주위의 형태를 파악한다. 이후 5대의 로봇들은 보다 넓은 자신의 활동범위를 보장하는 곳으로 이동하게 되므로 어떤 한 지점에서의 충돌 없이 탐색 공간을 돌아다닌다. 따라서 랜덤 탐색 보다 효율적으로 공간의 탐색을 수행할 수 있다. 실험 결과 평균 2대 정도의 로봇이 목표물을 찾아낼 수 있었다. 그림 10은 탐색에 성공한 로봇들을 (검은색 화살표로 표시된 로봇들) 보여준다.



그림 10. ABAM을 이용한 목표물 탐색.

Fig. 10. Five-robots are searching the object using ABAM.

5.2.3 Hexagonal area-based Q-learning을 이용한 탐색

마지막으로 hexagonal area-based Q-learning을 이용하

여 탐색을 시도하였다. Hexagonal area-based Q-learning은 로봇의 행동 결정상의 오류를 줄여주고, Q값의 학습을 통해 로봇의 효율적인 행동을 강화해 나갔다. 또한 시행을 거듭해 갈수록 로봇들은 주위의 환경에 대한 정보를 학습하였다. 결과적으로 평균 2대 이상의 로봇이 더 짧은 수행시간 내에 목표물을 발견하였다. 그림 11은 hexagonal area-based Q-learning을 이용한 탐색 과정을 보여준다.



그림 11. Hexagonal area-based Q-learning 을 이용한 목표물 탐색.

Fig. 11. Five-robots are searching the object using hexagonal area-based Q-learning.

5.2.4 실험결과

3개의 알고리즘을 적용한 실험의 결과를 그림 12에 나타내었다. 총 10회의 탐색 시도에서, 랜덤 탐색의 경우 2번째와 6번째 경우 각 1대의 로봇이 목표물을 찾아내었으나, 랜덤 탐색의 특성상 통계적인 의미를 부여하기는 어렵다. 다음으로 ABAM을 적용하였을 경우, 모든 시행동안 평균 1대 이상의 로봇이 목표물을 찾아내었다. 이것은 ABAM을 통해서도 탐색의 성능이 상당히 강화 될 수 있음을 나타낸다. 마지막으로 hexagonal area-based Q-learning을 통한 탐색의 결과는 주목할 만한 결과이다. 총 10회 시행에 평균 2대 이상의 로봇이 목표물 탐색에 성공하였다. 특히, 4번째, 6번째, 8번째, 9번째, 10번째 시행에서는 각 3대의 로봇이 목표물을 찾는데 성공하였다.

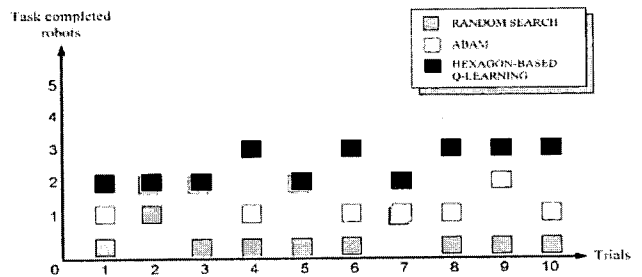


그림 12. 3개의 알고리즘을 적용하여 목표물 탐색을 수행한 결과.

Fig. 12. Experimental result with three different control methods.

6. 결론 및 향후과제

본 논문에서는 먼저 선행적 지식이 없고 장애물이 놓여있는 공간에서의 목표물 탐색 알고리즘으로 area-based action making process (ABAM)과 hexagonal area-based Q-learning을 제안하였다. 또한 실제 5대의 로봇을 통해 위에서 언급한 조건의 환경에서의 목표물 탐색을 수행하고 그 결과를 보였다. 실험의 결과를 통해 이 알고리즘이 위와 같은 환경에서의 목표물 탐색에 새로운 방법이 될 수 있음을

알 수 있다.

향후 과제로는 칠판, 로봇들의 협조 행동 구현을 위해 목표물 발견 후 목표물에 접근하는 알고리즘의 구현이 필요하다. 블록 쌓기나 물체 수송과 같은 작업을 위해서는 자연히 그립퍼(Gripper)의 장착이 필요하겠다. 둘째, 메인부와 서브 모듈간의 프로세스를 더 효율적으로 구성할 필요가 있다. 셋째, 다수 로봇에 의한 물체 수송, 대열을 갖춘 다수 로봇의 이동, object following 또는 path following 등의 로봇 기동에 관한 구현과 Fuzzy와 강화학습의 융합이나 TD(λ) 방법의 적용과 같은 심도 있는 알고리즘의 적용에 대한 연구가 뒤따라야 할 것이다.

참 고 문 헌

- [1] L. Parker, "Adaptive action selection for cooperative agent teams," *Proc. of 2nd Int. Conf. on Simulation of Adaptive Behavior*, pp. 442-450, 1992.
- [2] G. Ogasawara, T. Omata, and T. Sato, "Multiple movers using distributed, decision-theoretic control," *Proc. of Japan-USA Symp. on Flexible Automation*, vol. 1, pp. 623-630, 1992.
- [3] D. Ballard, *An Introduction to Natural Computation*, The MIT Press Cambridge, 1997.
- [4] J. Jang, C. Sun, and E. Mizutani, *Neuro-Fuzzy Soft Computing*, Prentice-Hall New Jersey, 1997.
- [5] W. Ashley, T. Balch, "Value-based observation with robot teams (VBORT) using probabilistic techniques," *Proc. of Int. Conf. on Advanced Robotics*, 2003.
- [6] W. Ashley, T. Balch, "Value-based observation with robot teams (VBORT) for dynamic targets," *Proc. of Int. Conf. on Intelligent Robots and Systems*, 2003.
- [7] D. Patterson and J. Hennessy, *Computer Organization and Design*, Morgan-Kaufmann Korea, 2005.
- [8] T. Mitchell, *Machine Learning*, McGraw-Hill Singapore, 1997.
- [9] C. Clausen and H. Wdchsler, "Quad-Q-learning," *IEEE Trans. on Neural Network*, vol. 11, pp. 279-294, 2000.
- [10] H-U. Yoon, S-H. Whang, D-W. Kim, and K-B Sim, "Strategy of cooperative behaviors of distributed autonomous robotic systems," *Proc. of 10th Int. Symp. on Artificial Life and Robotics*, pp. 151-154, 2005.
- [11] H-U. Yoon and K-B. Sim, "Hexagon-Based Q-Learning for Object Search with Multiple Robots," *Lecture Notes in Computer Science (LNCS)* published in Springer, vol. 3612, pp. 713-222, 2005.

저 자 소 개



윤한일(Han-UI Yoon)

2004년 : 중앙대학교
전자전기공학부 공학사
2004년~현재 : 동대학원
전자전기공학부 석사과정

관심분야 : Processor architecture, Distributed autonomous robotic system, Machine learning

Phone : +82-2-820-5319
E-mail : huyoon@wm.cau.ac.kr



심귀보(Kwee-Bo Sim)

1984년 : 중앙대학교 전자공학과 공학사
1986년 : 동대학원 전자공학과 공학석사
1990년 : The University of Tokyo 전자공학과 공학박사
1991년 ~ 현재 : 중앙대학교
전자전기공학부 교수
2003년 ~ 2004년 : 일본계측자동제어학회

(SICE) 이사
2000년 ~ 2004년 : 제어자동화시스템공학회 이사,
(현) 지능시스템연구회 회장
2003년 ~ 2004년 : 한국퍼지 및 지능시스템학회 부회장
2002년 ~ 현재 : 중앙대학교 산학연컨소시엄센터 센터장 및 기술이전센터 소장
2005년 ~ 현재 : 한국퍼지 및 지능시스템학회 수석부회장
2005년 : 제어자동화시스템공학회 Fellow 회원

관심분야 : 인공생명, 지능로봇, 지능시스템, 다개체시스템, 학습 및 적응알고리즘, 소프트 컴퓨팅(신경망, 퍼지, 진화연산), 인공면역시스템, 침입탐지시스템, 진화하드웨어, 인공두뇌, 지능형 홈 및 홈네트워크, 유비쿼터스 컴퓨팅 등

Phone : +82-2-820-5319
Fax : +82-2-817-0553
E-mail : kbsim@cau.ac.kr
Homepage URL : <http://alife.cau.ac.kr>