

# 군집 로봇의 협조 행동을 위한 강화 학습 기반의 진화 및 학습 알고리즘

## Reinforcement Learning Based Evolution and Learning Algorithm for Cooperative Behavior of Swarm Robot System

서상욱 · 김호덕 · 심귀보\*

Sang-Wook Seo, Ho-Duck Kim, and Kwee-Bo Sim

중앙대학교 전자전기공학부

### 요 약

군집 로봇시스템에서 개개의 로봇은 스스로 주위의 환경과 자신의 상태를 스스로 판단하여 행동하고, 필요에 따라서는 다른 로봇과 협조를 통하여 어떤 주어진 일을 수행할 수 있어야 한다. 따라서 개개의 로봇은 동적으로 변화하는 환경에 잘 적응할 수 있는 학습과 진화능력을 갖는 것이 필수적이다. 이를 위하여 본 논문에서는 새로운 Polygon 기반의 Q-learning 알고리즘과 분산유전알고리즘을 이용한 새로운 자율이동로봇의 행동학습 및 진화방법을 제안한다. 또한 개개의 로봇이 통신을 통하여 탐색체를 교환하는 분산유전알고리즘은 각기 다른 환경에서 학습한 우수한 탐색체로부터 자신의 능력을 향상시킨다. 특히 본 논문에서는 진화의 성능을 향상시키기 위하여 강화학습의 특성을 이용한 선택 교배방법을 채택하였다. 제안된 방법은 협조탐색 문제에 적용하여 컴퓨터 모의실험을 통하여 그 유효성을 검증한다.

### Abstract

In swarm robot systems, each robot must behaves by itself according to the its states and environments, and if necessary, must cooperates with other robots in order to carry out a given task. Therefore it is essential that each robot has both learning and evolution ability to adapt the dynamic environments. In this paper, the new polygon based Q-learning algorithm and distributed genetic algorithms are proposed for behavior learning and evolution of collective autonomous mobile robots. And by distributed genetic algorithm exchanging the chromosome acquired under different environments by communication each robot can improve its behavior ability. Specially, in order to improve the performance of evolution, selective crossover using the characteristic of reinforcement learning is adopted in this paper. we verify the effectiveness of the proposed method by applying it to cooperative search problem.

Key words : Behavior Learning, Evolution, Swarm Robot, Group Behavior, Cooperative strategy, Q-learning

### 1. 서 론

인간 사회의 모든 환경들은 인간이 사용하기에 가장 쉽고 편리하게 느낄 수 있도록 설계되어 가고 있다. 편안한 인간의 삶을 누리기 위하여 사회의 많은 부분에서 로봇이 이용되어져 가고 그 활용 범위도 점점 넓어져 가고 있다. 로봇들은 인간의 편리함뿐만 아니라 많은 위험한 분야에서 사람들 대신에 사용되어 지고 있다. 화재가 발생한 건물에서의 구조 활동이나 가스에 오염된 지역의 탐색 작업, 깊은 바다 속의 탐색, 극지방과 같은 곳에서의 기후 조사와 같은 부분에서 로봇이 사람을 대신하여 작업을 수행하고 있다. 뿐만 아니라

곤충의 집과 같은 사람이 직접 접근하기 힘든 곳의 탐색에서의 신뢰성과 이용가치가 높은 정보의 획득을 위해서 다수의 소형 로봇들이 보내지기도 한다[1].

근래에는 로봇을 이용한 사회 안전 분야에도 많은 연구가 진행되고 있다. 로봇을 이용한 경비, 탐색 등은 기존의 카메라나 적외선 센서와 같은 보안 시스템과 연동을 통해 이루어지고 있다. 이런 로봇 시스템의 사회적 인프라가 구축되어 가면서 기존의 한 대의 로봇을 가지고는 많은 업무를 수행할 수 없기 때문에 군집 로봇의 필요성이 증가되어져 가고 있다. 이런 군집 로봇 시스템을 제어하기 위해서 과거에는 중앙 집중식 제어를 많이 사용하고 있는데, 중앙 집중식 제어는 중앙에서 필요한 임무에 따라 모든 부분을 통제 할 수 있기 때문에 빠르고 정확한 제어가 가능하다는 장점을 가지고 있다. 그러나 제어해야 할 로봇 제어 시스템들이 거대화되고 복잡해짐에 따라서 로봇 제어 시스템의 유연성과 강인함이 점점 중요시 되어가고 있다. 이를 만족하기 위하여 여러 가지 알고리즘을 제안되어지고 있다. Jindong Tan은 ad hoc 무선 네트워크를 이용한 a fault tolerant 알고리즘을 이용해서 구조 제어 알고리즘을 제안하였으면[2], 이 제안을 이

접수일자 : 2007년 6월 22일

완료일자 : 2007년 9월 5일

\* 교신 저자

감사의 글 : 본 연구는 2007년 산업자원부의 [총괄과제 : 집단(群) 로봇 기술을 이용한 사회안전로봇 개발, 세부과제 : 로봇 통제 및 환경기술 개발] 연구비지원에 의해 수행되었습니다. 연구비지원에 감사드립니다.

용하여 Multi-Agent system의 더 발전된 모델을 제시하였다[4]. 또한 여러 많은 연구자들은 Neural Network, Fuzzy, Genetic Algorithm(GA), Soft-computing등을 사용하여 로봇의 제어에 이용하였다[6-9]. 군집로봇의 통신에도 많은 비중을 차지하면서 연구되어져 가고 있다[10-12]. 특히, 자율분산 제어시스템을 이용한 알고리즘들이 개발되어져 가고 있다. 분산 제어시스템을 이용해서 군집 로봇을 제어하는 여러 가지 방법들을 제안하였다. Vaithilingam Kumarathasan은 지리 탐지를 위해서 분산 제어를 이용한 로봇들의 협조행동을 이용하였다[3]. 그리고 Ahmad는 지능형 Multi-Agent Mobile-Robotics System을 위해서 좀 더 개발된 분산 제어를 사용하였다[13]. 자율 분산 제어 시스템은 인간 사회나 곤충의 군집체계, 생체 면역 시스템 등을 모델링하여 구성하는 모델 등이 있다. 자율분산 로봇 시스템은 이러한 자율 분산 제어시스템의 알고리즘을 이용하여 다개체(Multi-Agent) 로봇의 협조행동을 구성하고 있다. 군집 로봇간의 이동을 위해서 우리는 영역 기반 행동 알고리즘(ABAA)을 제안하였다. 이것은 곤충들의 행동에서 아이디어를 얻어서 제안된 알고리즘이다. 그리고 각각의 영역의 센싱을 위해서 Dodecagon(십이각형)을 기준으로 하였다. 그리고 그 알고리즘에서는 Q-learning을 같이 사용하였다. 다른 연구자들도 Q-learning Algorithm을 이용하여 군집로봇을 제어한 선례도 있다 [14-16]. 이 지능 학습은 군집로봇의 이동, 물체 회피, 물체 추적에 사용된다. 강화학습은 로봇의 환경의 여러 가지 영향을 지능 학습을 통해서 다음 행동을 결정함에 있어서 학습되어진 데이터에 의해서 행동하게 된다. 불확실한 영역의 탐색하는 동안 GA를 이용해서 reward를 받게 된다. 각각의 agent들은 계속해서 학습되어지고 다음 영역의 움직임 결정하게 된다. 본 연구에서는 Q-learning을 이용한 군집 행동과 물체추적에 초점을 맞추고 있다. Q-learning은 강화학습 기술로써 Q-learning이 불확실한 정보를 가진 Markovian action 문제를 해결하기 위해 가장 간단한 방법이기 때문에 이것을 사용하였다.

본 논문에서는 기존에 사용된 군집 제어방법과 Q-learning 학습 알고리즘을 사용하여 더 향상된 군집제어를 하였다. 특히, 군집제어에서 가장 중요한 중앙 집중식 제어와 분산 제어식 제어를 상황에 따라 동시에 사용하였다. 평상시에는 분산 제어식을 사용하여 로봇들을 제어하였고 이벤트 발생시 그 물체를 추적하기 위해서는 중앙 집중식을 사용하여 수행시간을 최적화하는데 초점을 두고 있다. 본 논문은 1장에서는 서론을 말하였고, 2장에서는 본 논문에 한 개의 로봇에 적용할 DBAM과 ABAM 방법과 Q-learning 알고리즘의 구조와 적용에 대해서 설명하였습니다. 3장에서는 자율이동 로봇의 군집제어에 두 가지 종류인 중앙 통제 방식과 분산 제어 방식에 대해서 설명하였다. 4장에서는 컴퓨터 시뮬레이션을 하여 실험과 결과를 보여주었다. 그리고 여러 가지 알고리즘에 따른 결과들을 비교하였고 마지막 5장에서는 결론 및 향후 과제에 대해 논의한다.

## 2. 자율이동로봇 주행을 위한 알고리즘

### 2.1. DBAM and ABAM

본 논문에서 한 대의 로봇이 다음 행동을 결정 하기위해 Distance-based action making process(DBAM)와 Area-based action making process(ABAM)를 사용하게 된다. DBAM 방식은 주위의 물체나 장애물과 로봇 사이의 거

리를 인식하여 가장 거리가 먼 곳으로 이동하게 된다. 만약에 거리가 같은 여러 곳을 인식한다면 로봇은 자신을 조금씩 회전하여서 거리를 유지하거나 더 멀어지는 곳이 있으면 그곳으로 로봇은 이동하게 된다. 그러나 ABAM 방식은 로봇이 주위의 인식을 위해서 주변을 지역으로 나누어서 인식을 하게하는 방식이다. 그리고 그 지역 중에서 가장 넓은 공간으로 이동하게 된다. 그림 1은 같은 상황에서 다른 행동을 보이는 것을 나타낸 것이다. 같은 곳으로 이동을 하는 경우도 있을 수 있게 된다. 그림 1에서 로봇이 12개의 센서를 가지고 있다는 가정에서 출발해 볼 수 있다. DBAM의 방식으로 보면 로봇은 d1, d9, d11 중에 한곳으로 이동하게 된다. 여기서 로봇은 반시계 방향으로 회전하면 d9나 d11의 길이가 점점 길어진다. 그래서 d9나 d11의 방향으로 이동하게 된다. 그러나 d9 방향으로 가게 되면 장애물 때문에 다시 방향을 잡고 이동을 해야 한다. 이러한 단점을 보완하기 위해서 ABAM을 사용하게 된다.

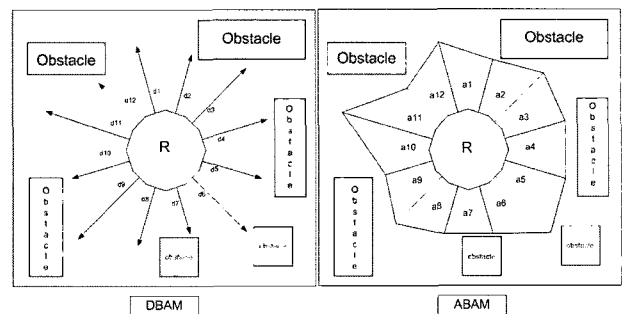


그림 1. 로봇의 행동을 위한 DBAM와 ABAM  
Fig. 1 DBMA and ABAM for robot behavior

그림 1의 ABAM을 보게 되면 로봇은 a10, a11 사이의 방향으로 이동하게 되는 것을 보게 된다. a10, a11의 영역의 넓이가 가장 넓기 때문에 넓은 부분으로 이동을 하게 된다. 그림 2는 로봇의 각각의 방법에 따른 충돌을 회피하면서 이동하는 것의 예를 보여준다. 이 로봇은 센서가 6개가 있다는 가정 하에서 보여준다. 그림 2에서 보여주는 시나리오에는 로봇 주위에 6개의 장애물이 있다. 먼저, DBAM에서는 남서쪽 방향으로 물체가 없다고 인식하고 이동하게 된다. 방향만을 의존해서 가기 때문에 두 개의 장애물에 사이에서 충돌하게 된다. 그러나 ABAM에서는 로봇이 두 장애물 사이의 거리를 계산해서 장애물 사이의 거리가 먼 북쪽 방향으로 이동하게 된다. 여기서 감지되지 않은 물체의 오류를 피하기 위해서 로봇은 좌우로 조금 회전을 해서 넓이를 측정하게 된다. 감지되지 않은 물체의 오류를 피하기 위해서 이 시나리오에 따르면 두 가지 경우 같은 위치에 있지만 다음 동작을 위해서 이동을 계산한다.

### 2.2. Q-Learning Algorithm

Q-learning은 강화학습으로 잘 알려진 알고리즘이다. 그리고 로봇이 효과적인 행동을 하기 위해서 보상의 개념을 이용해서 최적의 제어를 얻을 수 있다. 여기서 보상은 행동 후 보상을 하게 된다. Q-learning 알고리즘은 표 1에서 설명하는 것과 같다. 갱신될 Q 값은 다음의 식에 따라서 갱신된다.

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

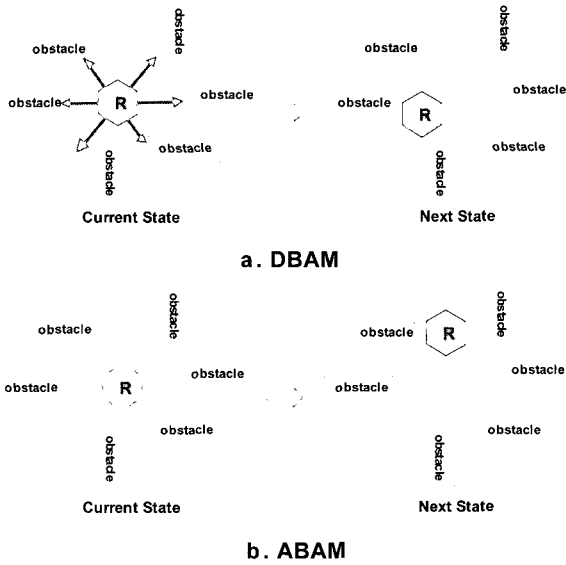


그림 2. DBAM와 ABAM에서 로봇의 충돌 회피  
Fig. 2. Collision avoidance of robot in DBAM and ABAM

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

<p>For each <math>s, a</math> initialize table entry <math>\hat{Q}(s, a)</math> zero</p> <p>observe the current state <math>s</math>.</p> <p>Continue to infinity.</p> <ul style="list-style-type: none"> <li>• Select action <math>a</math> and execute.</li> <li>• Receive immediate reward <math>r</math>.</li> <li>• Observe new state <math>s'</math>.</li> <li>• Select action.</li> <li>• Update table entry for <math>\hat{Q}(s, a)</math>.</li> <li>• <math>s \leftarrow s'</math>.</li> </ul>
---

그림 3은 Q-learning 알고리즘을 자세하게 설명한 그림이다. 로봇(R)은 다음 상태를 위해서 Q값이 높은 쪽으로 이동을 하게 된다. 그림 3의 예를 보면 현재 위치의 Q값은  $\hat{Q}(s_1, a_{right}) = 72$ ,  $a_{right}$ 는 로봇의 행동을 나타내는데 로봇은 오른쪽으로 다음 행동을 하게 된다. 로봇이 오른쪽으로 이동을 한다면  $r=0$ ,  $\gamma=0.9$ 의 미리 지정된 값으로 그 값들이 (2)식을 이용하여 갱신하게 된다.

$$\begin{aligned} \hat{Q}(s_1, a_{right}) &\leftarrow r + \gamma \max_{a_2} \hat{Q}(s_2, a_2) \\ &\leftarrow 0 + 0.9 \max_{a_2} \{63, 81, 100\} \\ &\leftarrow 90. \end{aligned} \quad (2)$$

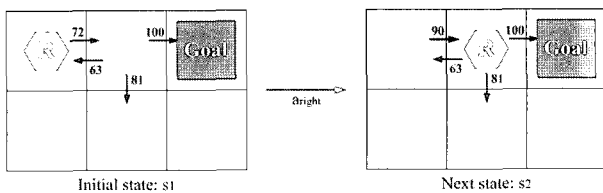


그림 3. Q-learning algorithm의 예  
Fig. 3. Example of Q-learning algorithm

### 2.3. 다각형을 기본으로 한 Q-learning의 적용

로봇 시스템을 위해서 특별한 Q-learning의 형태는 ABAM의 처리를 향상시키기 위해 적용하였다. 이 방법은 간단하면서도 쉽게 변경이 가능한 방법이다. Polygon-Based Q-learning이다. 다각형을 기본으로 만든 이유는 로봇의 센서의 개수에 따라서 같은 방법으로 적용이 가능하기 때문이다. 여기서는 6개의 센서를 가진 로봇을 예를 들어 설명하였다.

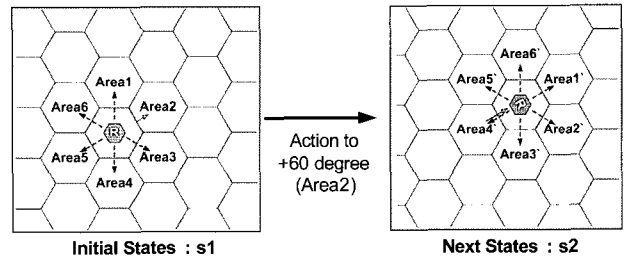


그림 4. Polygon-based Q-learning의 예  
Fig. 4. Example of Polygon-based Q-learning

그림 4는 다각형을 기본으로 한 Q-learning의 나타내었다. 여기서는 6개의 센서를 가진 로봇이라서 육각형을 기준으로 표현하였다. 현재 잘 알려져 있는 표준이 되는 Q-learning은 그림 3과 같은 사각형을 기본으로 되어있다. 본 논문에서는 기존의 사각형이 아닌 다각형의 여러 다른 모양을 기본으로 사용하는 Q-learning을 제안하였다. 많은 다른 종류의 로봇에 적용하기 위해서 다각형을 기본으로 하였다. 육각형, 팔각형, 십이각형 등 센서의 개수와 로봇의 모양에 다양하게 적용이 가능한 방법이다. 그리고 이렇게 여러 다각형을 이용하여서 다른 로봇과의 결합에 의해서 무한대로 확장이 가능한 것이다. 그림 4를 보면 육각형 모양의 6가지 방향으로 로봇이 이동할 수 있고 6개의 Q값을 입력을 가지고 있다. 그리고 활동 영역을 빠르게 응답할 수 있다. 그림 4를 가지고 설명해 보면, 로봇은 초기 상태를 가지고 있다. 그 로봇은 주위를 6개의 영역으로 나누고 60°의 범위를 가진다. 이동을 위해서 주위의 영역을 계산 후에 ABAM을 적용해서 가장 넓은 곳을 이동한다. 만약 Area2가 가장 넓은 영역이라면  $\hat{Q}(s_1, a_{+60^\circ})$ 의 값을 (1), (2)식을 이용해서 갱신하게 되는데, 그 갱신 과정은 식 (3)과 같다. 그리고 그림 4와 같이 로봇은 60°만큼 오른쪽으로 회전을 해서 이동전 Area2 영역이 Area1 영역으로 변하게 된다. 식에서  $s$ 는 가능한 상태를 나타내고,  $a$ 는 가능한 행동을 나타낸다. 그리고  $r$ 은 즉각적인 보상 값을 나타내고 초기 값으로 0이 주어지게 된다. 처음 상태에서 다음 상태로 이동한 후에는 즉각적인 보상은 변하게 된다. 그 값은 이동후의 가지는 모든 지역 값의 합과 이동전의 모든 지역 값들의 합의 차이를 보상받는 값으로 받게 되는데, 보상 값은 (4)식과 같이 계산된다.

$$\begin{aligned} \hat{Q}(s_1, a_{right}) &\leftarrow r + \gamma \max_{a_2} \hat{Q}(s_2, a'_\theta) \\ &\leftarrow 0 + 0.9 \max_{a_2} \{Area1', \dots, Area6'\} \\ &\leftarrow \gamma Area2'. \end{aligned} \quad (3)$$

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

각각  $Area_i \in s$ 와  $Area_j \in s'$ 의 영역 범위에 속해 있다. 다각형을 기반으로 한 Q-learning 알고리즘은 다음 표 2와 같다.

표 2. 다각형 기반 Q-learning 알고리즘  
Table 2. Polygon-based Q-learning algorithm

- For each  $s, a$  initialize the table entry  $\hat{Q}(s, a_\theta)$  zero  
calculate n-areas at the current state  $s$ .  
Do until the task is completed.
- Select action  $a_\theta$  to the widest area and execute.
  - Receive immediate reward  $r$ .
  - Observe new state  $s'$  if  $\hat{Q}(s', s'_\theta)$  is greater than or equal to  $\hat{Q}(s, a_\theta)$ .
  - Update table entry for  $\hat{Q}(s, a_\theta)$ .
  - $s \leftarrow s'$  otherwise, if  $\hat{Q}(s', s'_\theta)$  is far less than  $\hat{Q}(s, a_\theta)$ .
  - Move back to previous state.
  - $s \leftarrow s'$

### 3. 자율이동 로봇의 군집 제어

자율이동 로봇의 군집 제어는 행동 결정 주체에 따라 중앙통제 방식(Centralized Control)과 분산제어 방식(Distributed Control)으로 나누어서 하고 있다. 특히, 근래에는 분산 제어식의 방식을 둔 연구에 중점을 두고 이루어져 가고 있다. 그림 5에서와 같이 U-환경에서는 로봇들이 서로 같이 협력해서 같은 일을 하는 것보다는 서로 독립적인 일을 하면서 제어되어져 가고 있다. 본 논문에서는 평상시에는 분산제어 방식을 사용하고 모든 로봇이 통제에서 일시분란하게 일을 할 이벤트가 발생 시 중앙통제 방식에 따라 로봇들은 제어가 된다.

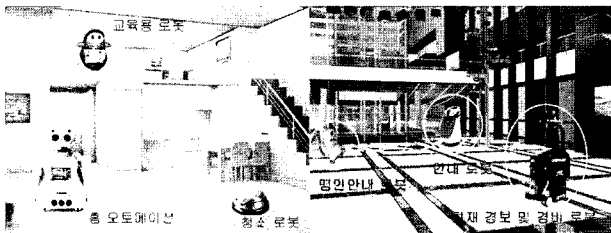


그림 5. U-환경에서의 coordinated multi-agent 로봇  
Fig. 5. Coordinated multi-agent robot in U-environments

그림 6은 군집 로봇의 협조 행동을 위한 흐름도를 나타내고, 그림 7은 자율 이동로봇의 군집 행동을 위한 학습과 진화의 개념을 나타낸다. 평상시에는 각각의 로봇이 분산 제어를 통해서 자율이동을 하고 자신의 임무를 수행한다. 이벤트 발생 시에는 한 개의 로봇 또는 중앙에서 모든 로봇에 명령을 전달하여 모든 로봇의 보다 빠르게 제어한다. 그림 7에서 로봇은 즉각적으로 반응할 수 있게 패턴을 미리 분류하고 학습한 상태에서 판단하고 행동을 하게 된다.

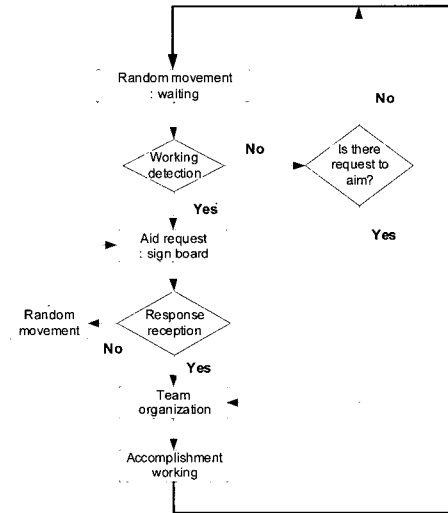


그림 6. 군집 로봇의 협조 행동을 위한 흐름도  
Fig. 6. Flowchart for cooperative behavior of swarm robot

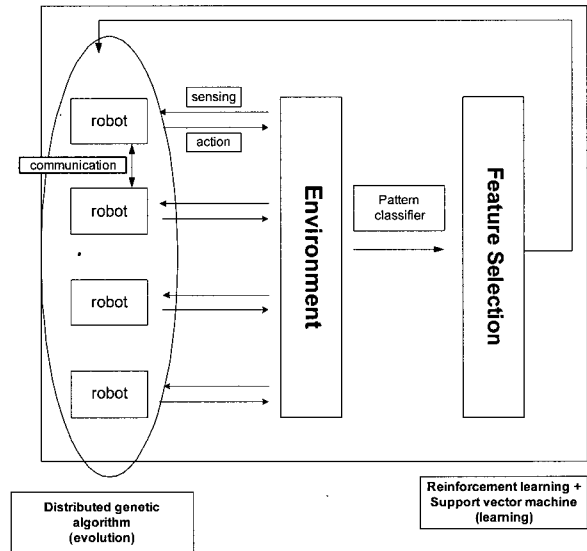


그림 7. 군집 로봇의 행동 학습과 진화  
Fig. 7. Behavior learning and evolution of swarm robot

#### 3.1 분산제어 방식과 중앙통제 방식

자율이동 로봇의 군집 제어는 분산제어 방식과 중앙통제 방식 이 두 가지를 같이 사용하게 된다. 먼저 분산 제어 방식은 각각의 로봇의 개별 개체들이 모두 독립적으로 서로 간의 임무를 수행하는 방식이다. 다양한 정보를 동시에 취득하고 교류하는데 유리하다. 그러나 단체 행동의 최적화 및 효율성에 취약한 단점을 가지고 있다. 분산제어 방법에는 중력장(그림 8a), 로드맵(그림 8b), 분산주행(그림 8c), 행위기반, 셀 분해 방식 등이 일반적으로 사용되는데, 그림 8은 대표적인 분산제어 방식의 예를 나타낸 것이다.

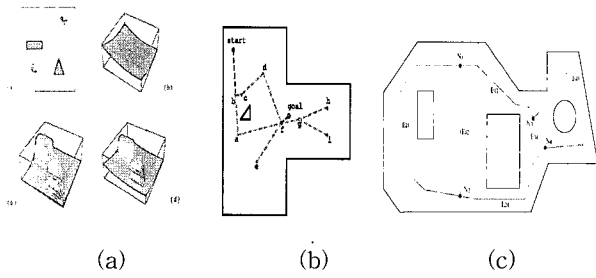


그림 8. 군집 로봇의 분산제어 방법의 예  
 Fig. 8. Representative examples of distributed control method for swarm robot

한편 중앙 통제 방식은 중앙제어 컴퓨터에서 모든 로봇을 통제하고 개체간의 간섭과 충돌 없이 일사 분란한 통제에 유리하다. 그러나 분산제어의 장점인 다양한 정보수집 및 자율성에는 취약한 약점을 가지고 있다. 그림 9와 같이 군집 로봇의 대형 주행 기법에서 사용될 수 있다.

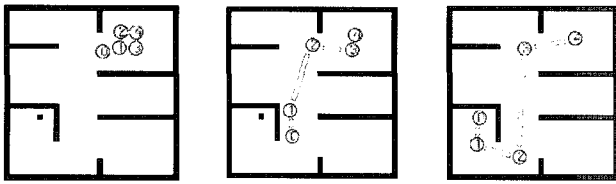


그림 9. 군집 로봇의 대형 주행  
 Fig. 9. Follower the leader of swarm robot

본 논문에서 위에서 언급한 분산 제어 방식과 중앙 통제 방식을 상황에 따라 사용하게 된다. 평상시는 다양한 정보수집 및 자율성에 유리한 분산 제어 방식을 사용하고, 침입자 발견과 같은 이벤트 발생 시 이동할 때는 중앙 통제 방식을 사용한다. 본 논문에서는 컴퓨터가 통제 하는 방식보다는 한 대의 로봇이 이끌어가는 방식으로 실험하였다. 그리고 침입자에 접근해서는 각각의 입구에 맞추어서 포위를 하게 된다. 이때에는 군집 로봇의 제어 방식을 사용하였다.

4. 컴퓨터 모의실험 및 결과

본 논문에서 제안한 알고리즘을 군집 로봇의 두 가지 협조 행동에 대한 모의실험을 수행하였다. 첫 번째 모의실험은 협조 행동을 통해서 모든 로봇이 한곳에 모이는 실험이며, 두 번째 모의실험은 모든 로봇이 목표물을 추적해 가는 실험이다.

그림 10은 10대의 로봇이 왼쪽 아래 부분으로 모이는 모의실험을 한 것이다. 본 모의실험에서는 로봇이 서로 협조해서 움직이다가 모든 로봇이 최종 목표 지점에 도달하면 하나로 뭉쳐지고 모의실험이 끝난다. 모의실험 과정은 그림의 A→B→C→D의 순서이다.

그림 11은 5대의 로봇이 한 개의 표적을 찾아다니는 것을 나타내는 모의실험이다. 이 모의실험에서는 로봇이 주변 환경의 지도를 가지고 있다고 가정하였다.

목표물은 창의 왼쪽 위에 있는 녹색으로 표시된 것이고, 빨간색의 5대 로봇이 주변을 돌아다니면서 녹색의 목표물을 추적하게 된다. 이 모의실험에서는 움직이는 목표물에 대한

모의실험도 수행하였는데, 이는 이동 물체의 추적에 대한 성능을 확인하기 위해서였다. 목표물 추적이 끝나게 되면 화면에 “추적 완료”라는 메시지를 보이고 모의실험이 끝난다. 모의실험 과정은 그림의 A→B→C→D의 순서이다.

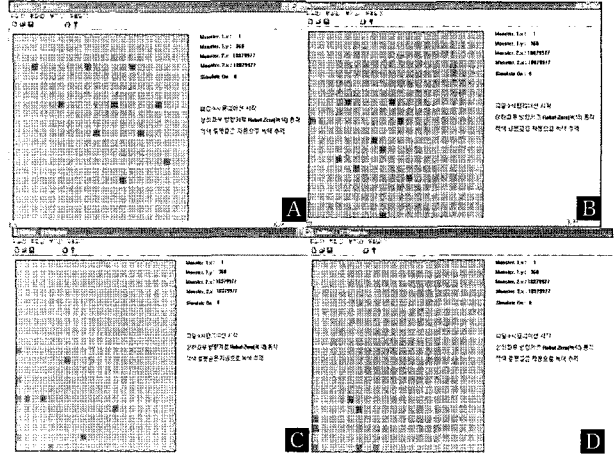


그림 10. 군집 로봇의 협조 행동 모의실험  
 Fig. 10. Simulation of cooperative behavior of swarm robot

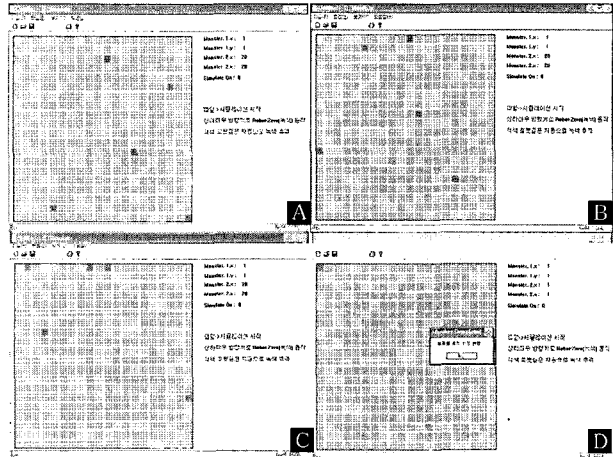


그림 11. 군집 로봇의 목표물 추적 모의실험  
 Fig. 11. Simulation of target tracking of swarm robot

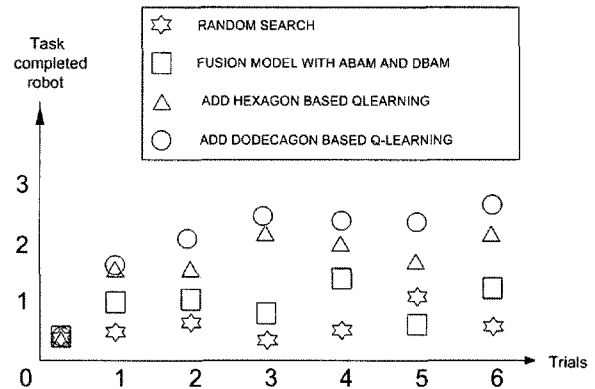


그림 12. 군집 로봇의 협조 행동을 위한 모의실험 결과  
 Fig. 12. Simulation results of swarm robot for cooperative behavior

군집 로봇의 협조 행동을 위한 모의실험 결과를 그림 12에 나타내었다. 이 실험 결과는 비교를 위해서 랜덤 서치, DBAM과 ABAM 방법, 기존의 Q-learning과 본 논문에서 제안한 Polygon-based O-learning 방법 등을 군집 로봇이 목표물을 추적하는 실험을 통해서 얻어진 결과이다. 그림을 보면 알 수 있듯이 본 논문에서 제안한 방법이 우수함을 알 수 있다.

### 5. 결론 및 향후과제

본 논문에서는 군집 로봇의 이동에서 ABAM 방법과 Polygon-Based Q-learning 알고리즘을 제안하였다. 군집 제어에 있어서는 중앙 통제 제어와 자율 분산 제어를 사용해서 로봇의 군집행동을 하였는데, 본 논문에서는 자율 분산 제어를 통해서 목표물을 추적하는데 중점을 두고 실험을 하였다. 그리고 제안한 군집 로봇의 목표물 추적 알고리즘은 모의실험을 통하여 그 유효성을 확인하였다. 모의실험 결과에서 알 수 있듯이 모든 로봇이 목표물 추적이 성공하였다. 특히 주목할 것은 미지의 환경에서 군집 로봇의 협조 행동을 위한 자율 주행은 기존의 격자 기반의 Q-learning 알고리즘을 사용했을 때보다 여러 가지 다각형을 사용함으로써 더 성능이 향상되는 것을 알 수 있었다. 이는 로봇이 더 많은 센서를 사용해서 주위 환경을 잘 인식하면 성능이 향상됨을 의미한다. 앞으로 군집 로봇의 성능 향상을 위해서는 로봇간의 통신 방법, 환경에 빠르게 적응하기 위한 학습과 진화 방법 등에 대한 연구가 필요하다.

### 참 고 문 헌

[1] 윤한얼, 심귀보, "다수 로봇의 목표물 탐색을 위한 Area-Based Q-learning 알고리즘", *한국퍼지 및 지능시스템학회 논문지*, 제15권, 제4호, pp. 406-411, 2005.

[2] Jindong Tan, Ning Xi, Weihua Sheng and Jizhong Xiao, "Modeling Multiple Robot Systems for Area Coverage and Cooperation", *Proc. of the 2004 IEEE International Conference on Robotics & Automation*, vol. 3, pp. 2568-2573, New Orleans, LA, April, 2004.

[3] Vaithilingam Kumarathan, Thrishantha Nanayakkara, "Intelligent Collaboration among Robotic Agents for Landmine Detection", *Proc. of the Annual Sessions of the Sri Lanka Association for Artificial Intelligence*, 2005.

[4] Alessabdro de Luna Almeida, Samir Akinine, Jean-Pierre Briot, Jacques Malenfant, "Plan-Based Replication for Fault-Tolerant Multi-Agent Systems", *Proc. of the 20th International Parallel and Distributed Processing Symposium*, 2006.

[5] Izzet Can Envarli, Julie A. Adams, "Task Lists for Human-Multiple Robot Interaction", *Proc. of the IEEE International Workshop on Robots and Human Interaction Communication*, pp. 119-124, Aug. 2005.

[6] Alain Cardon, Thierry Galinho, Jean-Philippe Vacher, "Genetic algorithm using multi-objectives

in a multi-agent system", *Proc. of Robotics and Autonomous System*, pp. 179-190, 2000.

[7] Xiaojiang Zhang, "Fuzzy control system for a mobile robot collision avoidance", *Pro. of the IEEE International Conference on Industrial Technology*, pp. 125-128, 1994.

[8] Ding Yingying, He Yan, Jiang Jing-Ping, "Self-Organizing Multi-robot System Based on Personality Evolution", *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*, 2002.

[9] Prasanna Sridhar, Shahab Sheikh-Bahaei, Shan Xia, Mo Jamshidi, "Multi agent Simulation using Discrete Event and Soft-computing Methodologies", *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 1, pp 1004-1012, Dec, 2003.

[10] Thomas W. Dunbar, Joel M. Esposito, "Artificial Potential Field Controllers for Robust Communications in a Network of Swarm Robots", *Proc. of the Thirty-Seventh Southeastern Symposium*, pp. 401-405, March. 2005.

[11] Mary Berna-Koes, Illah Nourbakhsh, Katia Sycara, "Communication Efficiency in Multi-Agent Systems", *Proc. of the IEEE International Conference on Robotics & Automation*, pp. 2129-2134, April. 2004.

[12] Chris A. C. Parker, Hong Zhang, "A Practical Implementation of Random Peer-to-Peer Communication for a Multiple-Robot System", *Proc. of the IEEE International Conference on Robotics and Automation*, pp. 3730-3735, April. 2007.

[13] Mohd Ridzuan Ahmad, Shamsudin H.M. Amin, Rosbi Mamat, "Development of Decentralized Based Reactive Control Strategy for Intelligent Multi-Agent Mobile Robotics System", *Proc. of the Seventh International Conference on Control, Automation, Robotics And Vision*, pp. 220-227, Dec. 2002.

[14] Ou Haitao, Zhang Weidong, Zhang Wenyuan, Xu Xiaoming, "A novel multi-agent Q-learning algorithm in cooperative multi-agent system", *Proc. of the 3rd World Congress on Intelligent Control and Automation*, pp. 272-276, 2000.

[15] Jing Huang, Bo Yang, Da-You Liu, "A Distributed Q-learning Algorithm for Multi-Agent Team Coordination", *Proc. of the Fourth International Conference on Machine Learning and Cybernetics*, pp. 108-113, August. 2005.

[16] Tong Zhou, Bing-Rong Hong, Chao-Xia Shi, Hong-Yu Zhou, "Cooperative Behavior Acquisition Based Modular Q-learning in Multi-Agent System", *Proc. of the Fourth International Conference on Machine Learning and Cybernetics*, pp 205-210, August. 2005.

저 자 소 개



서상욱(Sang-Wook Seo)  
2007년~현재 : 중앙대학교 대학원  
전자전기공학부 석사과정  
재학 중

[제17권 4호 (2007년 8월호) 참조]



심귀보(Kwee-Bo Sim)  
1991년 ~ 현재 : 중앙대학교  
전자전기공학부 교수

[제17권 4호 (2007년 8월호) 참조]



김호덕(Ho-Duck Kim)  
2006년~현재 : 중앙대학교 대학원  
전자전기공학부 석사과정  
재학 중

[제17권 4호 (2007년 8월호) 참조]