논문 2008-03-16

다각형 기반의 Q-Learning과 Cascade SVM을 이용한 군집로봇의 목표물 추적 알고리즘

(Object Tracking Algorithm of Swarm Robot System for using Polygon Based Q-Learning and Cascade SVM)

서 상 욱, 양 현 창, 심 귀 보*
(Sang-Wook Seo, Hyung-Chang Yang, Kwee-Bo Sim)

Abstract : This paper presents the polygon-based Q-leaning and Cascade Support Vector Machine algorithm for object search with multiple robots. We organized an experimental environment with ten mobile robots, twenty five obstacles, and an object, and then we sent the robots to a hallway, where some obstacles were lying about, to search for a hidden object. In experiment, we used four different control methods: a random search, a fusion model with Distance-based action making (DBAM) and Area-based action making (ABAM) process to determine the next action of the robots, and hexagon-based Q-learning and dodecagon-based Q-learning and Cascade SVM to enhance the fusion model with DBAM and ABAM process.

Keywords : Cascade SVM, Polygon, Q-Learning, DBAM, ABAM

1. 서론

최근 로봇 응용분야의 확장으로 다수의 로봇으로 구성된 시스템에 대한 연구가 많이 이루어지고 있는데, 특히 자율분산성을 가지는 군집 로봇시스템에 관한 연구가 관심을 모으고 있다. 이러한 군집 로봇시스템은 중앙관리형 시스템에 비하여 몇 가지 특징을 가지는데, 첫째로 각각의 로봇은 주변의 환경이나 물체 등을 인식하여 자신의 행동을 독립적으로 결정하며 주어진 작업을 잘 수행하기 위하여 다른 로봇과 협동할 수 있다. 둘째로는 강건성 (robustness)과 유연성 (flexibility)을 가지고 있다는 것이다. 몇 대의 로봇이 고장이 나더라도 시스템의 정상적인 동작에 영향을 주지 않으며, 주어진 일에 대하여 오직 로봇의 행동 규칙만 바꾸어 줌으로써 여러 가지 작업에 적용할 수 있다. 마지막으로는 시스템의 크기가 커지더라도 개개의 로봇은 자신의 주변상황에 따라 자신의 일을 판단하여 결정하므로 시스템의 복잡도가 증가하지 않는다 [1].

군집 로봇시스템을 제어하기 위해서 과거에는 중앙 집중식 제어를 많이 사용 하였는데, 중앙 집중식 제어는 중앙에서 필요한 임무에 대한 권한을 모두 가지고 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 있는 상황에서 모든 권한이 한 곳에 집중되어 진다는 것이 단점으로 작용될 수도 있다. 군집 로봇시스템의 제어 방법에 관해서 Parker는 다수 로봇의 작업 수행을 위해 heuristic 형태의 알고리즘을 제안하였고 [2], Ogasawara는 다수의 로봇을 이용해 커다란 물체를 수송하기위해 자율 분산 로봇 제어 방식을 이용하였다[2]. 본 논문에서는 군집 로봇시스템을 제어하기 위한 방법으로 분산식 제어 방법을 가정하였고, 다수의 로봇이 어떤 작업을 수행함에 있어 서로간의 충돌을 피하고, 자신만의 고유한 영역을 탐색하여 다음 행동을 진행

* 교신저자(Corresponding Author)
논문접수 : 2008. 4. 23., 채택확정 : 2008. 6. 13.
심귀보 : 중앙대학교 전자전기공학부
서상욱, 양현창 : 중앙대학교 전자전기공학부
※ 본 논문은 지식경제부의 2008년도 성장동력기술개발사업인 「집단 로봇 기술을 이용한 사회안전로봇 개발(세부과제: 로봇통제 및 환경기술개발)」에 의해 수행되었음.

하기 위한 방법으로 Distance Based Action making (DBAM)과 Area Based Action Making (ABAM)의 융합 모델을 사용하였다.

강화 학습은 agent로 하여금 주변 환경의 탐색을 통해 능동적으로 환경에 대한 행동을 결정하도록 한다. 보상 값이 존재하는 어떤 불확실한 영역을 탐색하는 동안 agent는 연속적인 상태 공간을 따라 적절한 보상 값을 전달함으로써, 임의의 상태에 대해 어떠한 행동을 취해야 할지를 학습하게 된다 [4]. 강화 학습을 구현하기위한 많은 방법 중, 본 논문에서는 Cascade Support Vector Machine과 Q-learning을 이용하였다 [5,6]. 또한 이 임의의 상태 공간은 실제로 물리적인 공간으로 간주될 수 있다. 그리고 결과로 나타난 Q-값에 대하여 Cascade SVM을 사용함으로써 로봇의 행동에 대한 오차를 줄일 수 있었다 [7]. 본 논문에서는 DBAM과 ABAM의 융합 모델을 강화시키기 위해 12각형 기반의 Q-learning과 Cascade SVM algorithm을 적용한다.

본 논문의 2장에서는 로봇의 행동 결정 과정에서 DBAM과 ABAM의 융합 모델에 대해 나타낸다. 3장에서는 다각형 기반의 Q-learning 알고리즘과 Cascade SVM에 대하여 논하며, 4장에서는 무작위, DBAM과 AMAB의 융합 모델, 6각형 기반의 Q-learning 알고리즘, 12각형 기반의 Q-learning과 Cascade SVM 알고리즘들을 이용하여 목표물 탐색의 시뮬레이션 및 결과를 보인다. 마지막으로 5장에서는 결론 및 향후 과제에 대해 논한다.

Ⅱ. 로봇의 행동 결정 과정

1. DBAM and ABAM

Distance-based action making (DBAM) 과 Area-based action making (ABAM) 방법은 로봇이 다음 행동을 결정하는데 있어서 사용되는 방법이다. DBAM 방법은 로봇이 주위의 환경을 거리로 인식하는 방법으로 로봇과 물체 사이의 거리에 의해서 행동을 결정한다. 반면 ABAM 방법은 로봇이 자신 주변의 환경을 둘 사이의 거리가 아닌 자신 주변의 면적을 계산하여 얻어진 정보로부터 다음 행동을 결정하는 방법이다. ABAM 방법의 핵심은 로봇으로 하여금 자신 주위의 불확실성을 줄여 나가도록 한다는데 있는데, 결국 ABAM 방법은 행동 기반 방향 전환 (behavior-based direction change) 방식과 많은 유사점을 가지고 있다고 할 수 있다 [8,9]. 그림 1은 DBAM과 ABAM 방법이 행동을 선택하는 기준을 나타내고 그림 2는 같은 환경 아래 로봇이 존재한다고 가정하였을 때 DBAM과 ABAM 방법이 어떻게 다음 행동을 선택하는지에 대한 방법을 예를 들어 설명하고 있다. 그림 2에서 보면 같은 상황이라도 행동 결정 방식에 따라서 차이가 있는데, 우선 DBAM 방법의 경우에는 로봇 주위의 상황을 장애물과 로봇사이의 거리만으로 결정을 하기 때문에 로봇이 취할 수 있는 다음 행동 자체가 많지만, ABAM 방법의 경우에는 넓이를 기반으로 다음 행동을 결정하기 때문에 로봇이 다음 행동을 하는데 있어서 제한적일 수는 있지만, 부정확한 행동을 조기에 차단할 수 있다는 장점이 있다.

2. DBAM and ABAM의 융합모델

본 논문에서는 로봇이 다음 행동을 결정하는데 있어서 DBAM과 ABAM 방법을 융합한 모델을 사용한다. DBAM 방법을 통해서 로봇으로부터 가장 거리가 먼 방향을 선택하고, ABAM 방법으로 주위 환경에서 가장 넓이가 큰 공간을 선택하게 된다. 단순히 거리만으로 다음 행동을 결정하는 DBAM 방법은 계산 량을 줄일 수 있다는 장점이 있지만, 올바르지 못한 행동을 선택할 확률이 높은 단점이 있다. 반면 행동을 선택할 때 단순히 넓이만을 고려하는 ABAM 방법은 올바른 행동을 선택할 확률이 있다는 장점이 있으나, 계산양이 많다는 단점이 있다.

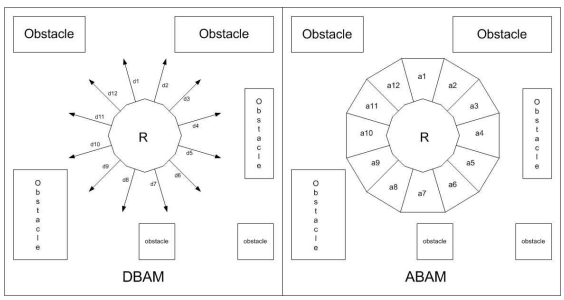

그림 1. DBAM과 ABAM의 행동 선택 기준.
Fig. 1. Action selection basis of DBAM and ABAM

DBAM과 ABAM 방법의 융합 모델은 행동을 선택할 때 거리가 가장 멀고, 넓이가 가장 큰 곳을 선택한다. 기본적으로 로봇과 물체와의 거리가 로봇의 반경을 12등분 하였을 때 장애물이 차지하는

넓이를 뺀 것보다 더 우선권이 있다고 가정하였기 때문에 본 논문에서 선택한 방법은 로봇으로부터 장애물과의 거리가 가장 먼 방향을 선택하고, 선택되어진 방향 중에서도 장애물이 차지하는 넓이가 가장 작은 공간을 선택하는 방법을 선택하였다. 그림 3은 DBAM과 ABAM 방법의 융합 모델에 관한 그림이다.

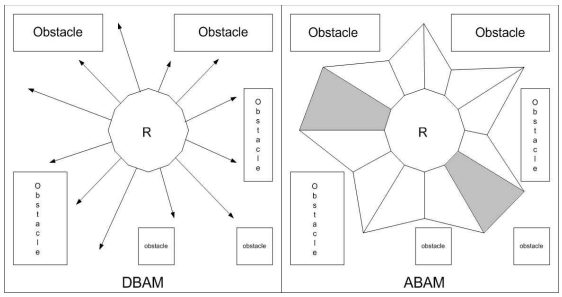

그림 2. DBAM과 ABAM의 행동 선택의 예.
Fig. 2. An illustrative example of action selection with DBAM and ABAM

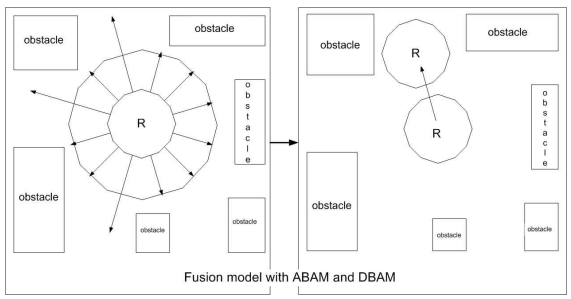

그림 3. DBAM과 ABAM의 융합모델
Fig. 3. Fusion model with DBAM and ABAM

Ⅲ. 다각형 기반의 Q-learning과 Cascade SVM 알고리즘

1. Q-learning

Q-learning은 강화학습으로 잘 알려진 알고리즘이다. 그리고 로봇이 효과적인 행동을 하기 위해서 보상의 개념을 이용해서 최적의 제어를 얻을 수 있다 [10]. 여기서 보상은 행동 후 보상을 하게 된다. Q-learning 알고리즘은 표 1에서 설명하는 것과 같다. 여기서 s는 상태를, a는 행동을, r은 보상값을, γ는 Q-함수 값의 조정을 위한 계수(discount factor)이다. 또한 갱신될 Q 값은 다음의 식 (1)에 따라서 갱신된다.

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

그림 4는 Q-learning에 대한 실제적인 예를 보여준다. 각각의 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 $\hat{Q}(s_1, a_{right}) = 72$와 같다.

표 1. 모델링을 위한 초기 파라미터들.
Table 1. Initial parameters for modeling

$\hat{Q}(s,a)$ 테이블의 모든 s와 a를 0으로 초기화 한다.
다음의 과정을 에피소드의 반복한다.
1. 현재 상태 s를 얻는다.
2. 현재 상태 s에 대하여 행동 a를 선택하고 행동 a를 수행한다.
3. 행동 a에 대한 즉각적인 보상 값 r을 얻는다.
4. 새로운 상태 s'을 얻는다.
5. 새로운 상태 s'에 대하여 모든 a'를 고려하여 행동 a'을 선택한다.
6. $\hat{Q}(s,a)$ 테이블에 값을 $s \leftarrow s'$, $a \leftarrow a'$로 업데이트 한다.

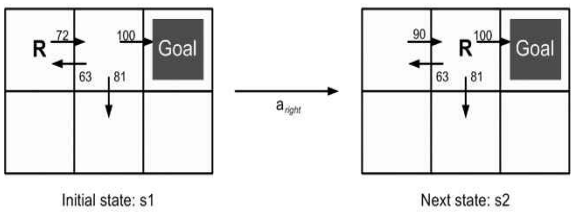

그림 4. Q-learning의 예
Fig. 4. An illustrative example of Q-learning

초기 상태에서 만약 로봇이 오른쪽으로 행동을 취한다면, 업데이트 되는 Q값은 $r=0$, $\gamma=0.9$을 초기 값으로 할 때

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

이 된다.

2. 6각형 기반의 Q-learning

첫 번째 방법으로 면적 기반 행동 결정을 강화시킨 Q-learning 알고리즘을 사용하였다 [11]. 로봇이 6개의 센서를 가지고 있다고 가정하면 기반이 되는 면적의 모양은 육각형이 된다. 따라서 어떤 임의의 상태의 로봇은 6방향으로 행동을 취할 수 있고, 6개의 Q값을 갖게 된다. 그림 5에 Q-learning 적용의 간단한 예를 나타내었다.

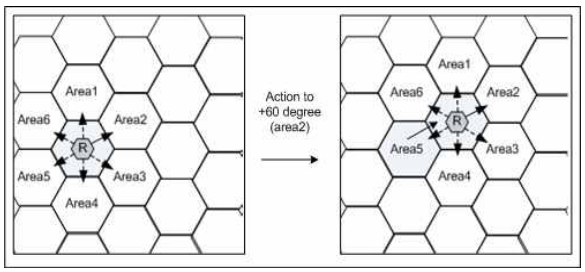

그림 5. 6각형 면적 기반 Q-learning의 예
Fig. 5. An illustrative example of Hexagon based Q-learning

만약 초기 상태에 서 로봇이 +60°방향으로 행동을 결정하였고, 상태 천이 후 Area2가 가장 넓은 영역을 보장하였다고 가정하면, 이에 대한 Q값은 $\hat{Q}(s_1, a_{+60})$로 표현될 수 있고, 그 값은 0이 초기 상태에서의 보상 값이라고 했을 때,

$$\begin{aligned}\hat{Q}(s_1, a_{+60}) &\leftarrow r + \gamma \max_{a'} \hat{Q}(s_2, a') \\ &\leftarrow 0 + \gamma \max(Area1, Are2, \cdots, Area6) \\ &\leftarrow \gamma Area2\end{aligned} \tag{3}$$

과 같이 된다. 초기 상태 이후의 보상 값은 다음과 같이 결정된다.

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

결과적으로 로봇은 이 Q값을 학습함으로써 진행 경로를 결정할 수 있다 [12].

3. 12각형 기반 Q-learning과 Cascade SVM 알고리즘

기존의 6각형 기반의 Q-learning은 면적 기반 행동 결정을 강화하기 위해 제안된 방법인데, 기반이 되는 단위 면적은 6각형으로 하였다. 그 이유로는 단위 면적의 조합에 따라 무한히 확장될 수 있는 다각형은 4각형을 제외하면 6각형뿐이기 때문이다. 또한 기존의 실험은 6개의 적외선 센서를 가진 로봇을 바탕으로 제안된 알고리즘이다. 반면 본 논문에서 제안한 12각형 기반 Q-learning과 Cascade SVM 알고리즘은 면적과 거리를 동시에 고려하여 다음 행동을 결정하게 된다. 즉 12개의 초음파 센서를 이용하여 12방향으로 물체를 측정하고 로봇 주위의 넓이를 12등분하여 인식한 후 각 넓이에서 물체의 넓이를 뺀 나머지 부분의 넓이가 가장 넓고, 물체와 로봇사이의 거리가 가장 긴 방향으로 행동을 취한다.

본 알고리즘은 기존의 Q-learning 알고리즘과 가장 큰 차이점은 업데이트를 하는 방법이다. 기존의 Q-learning 알고리즘에서는 이전의 Q 값과 비교를 해서 최소화시킬 수 있는 방향으로 이전의 상태를 좋은 방향으로 학습해 나가는 방법이다. 하지만 본 연구에서 제안된 12각형 기반의 Q-learning 과 Cascade SVM 알고리즘은 이전의 Q 값에 의존하지 않고 그 상황에서 최적의 행동을 결정하게 된다. Q 값의 갱신은 다음의 식 (5)로 이루어진다.

$$\begin{aligned}\hat{Q}(s,a) &\leftarrow r + \gamma \max_{s'} \hat{Q}(s', a) \\ &\leftarrow r + \gamma \max_{s'} (S_{area1} - S_{obstacle}) l_{obstacle}, \cdots, \\ &\qquad (S_{area12} - S_{obstacle}) l_{obstacle}\end{aligned} \tag{5}$$

$S_{areaN} - S_{obstacle}$는 로봇이 포함되어 있는 각 영역에서 장애물의 넓이를 뺀 값을 의미하고, $l_{obstacle}$은 로봇과 장애물 사이의 거리를 의미한다.

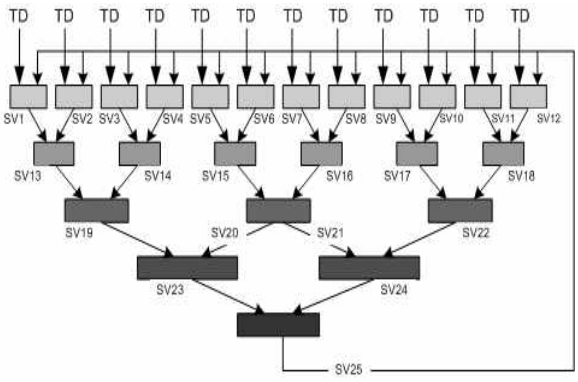

그림 7. 종속적인 SVM의 구조
Fig. 7. Structure of cascade SVM

여기서 $S_{areaN} - S_{obstacle}$이 가질 수 있는 12가지의 값과 $l_{obstacle}$이 가질 수 있는 12가지의 값이 각 Training Data (TD)가 되어서 입력 값으로 들어가

고 25개의 SVM 분류기로부터 얻어진 25개의 Support Vector를 통해서 최종적으로 가장 적절한 값을 결정하게 된다. 그림 7은 본 논문에서 사용한 Cascade SVM의 구조이다.

전체 training data 집합이 12개의 세부집합으로 이루어졌다고 가정하면 Cascade SVM 알고리즘은 다음과 같은 순서를 통해서 진행된다.

Step1. SVM 분류기를 생성하기 위해 $TD_1, TD_2, ..., TD_{12}$을 학습시킨다. 첫 번째 층에서 생성된 support vectors의 집합을 $SV_1, SV_2, ..., SV_{12}$이라고 한다.

Step2. $SV_1, SV_2, ..., SV_{12}$을 결합시키고 결합된 데이터 집합을 학습시켜 $SV_{13}, SV_2, ..., SV_{18}$을 생성한다. 3번째 층과 4번째 층에서도 동일한 과정을 거쳐서 $SV_N (N = 1,2,...,25)$을 생성한다.

Step3. $TD_2, TD_3, TD_4, ..., TD_{12}$에 SV_1을 추가하고 $TD_1, TD_3, TD_4, ..., TD_{12}$에 SV_2을 추가한 후 step2와 동일한 방법으로 학습시킨다. 여기서 생성된 support vectors의 집합을 $SV_1', SV_2', SV_3', ..., SV_{25}'$이라고 한다.

Step4. 만약 SV_1와 SV_1', SV_2와 SV_2', SV_3와 SV_3',..., 그리고 SV_{25}와 SV_{25}'이 차이점이 없거나 차이점이 있는 집합의 각 요소들이 고정된 값이 나오거나, 혹은 그 차이가 주어진 값 ϵ보다 작을 때는 최종적으로 나오는 결과 값은 SV_{25}'가 되고 그때 SVM 함수를 얻을 수 가 있다.

Step5. 만약 SV_1와 SV_1', SV_2와 SV_2', SV_3와 SV_3',..., 그리고 SV_{25}와 SV_{25}'이 위의 조건에 해당되지 않으면 step3으로 돌아간다.

Cascade SVM은 기존의 SVM에 비해 모든 SVM이 전체 Data 집합을 다루는 것이 아니라 각각의 SVM이 한 가지씩의 Traning Data를 다룸으로써 학습 속도가 향상되고, 12가지 상황 전부에 대한 분류를 할 수 있다는 장점이 있다.

결국 본 논문에서 제안한 12각형 기반 Q-learning과 Cascade SVM 알고리즘에서 Cascade SVM의 역할은 피드백을 얼마나 시킬 것인지 혹은 Support Vector를 어떻게 형성할 것인지에 따란 실험 결과에 많은 영향을 미칠 수 있다.

Ⅳ. 시뮬레이션 및 결과 고찰

본 연구에서 제안한 알고리즘의 유효성을 보이기 위해서 본 논문에서는 4개의 알고리즘을 이용하여 모의실험을 수행하였고, 그 결과를 그림 8, 9, 10, 11에 나타내었다. 그래프에서 가로축은 시행횟수를 나타내고, 세로축은 목표물을 찾은 로봇의 수를 나타낸다. 실험 환경에는 10대의 로봇과 25개의 장애물, 그리고 1개의 목표물이 있으며, 작업공간은 100×100 (로봇의 크기는 1×1, 목표물의 크기 1×1)으로 가정하였고, 탐색시간은 1회당 총60sec로 총 100번에 걸쳐서 진행하였다. 시뮬레이션의 map 특성상 랜덤으로 주어졌기 때문에 각각의 실험을 30번 정도 반복을 하여 그 결과 값을 평균하여 그래프로 나타내었는데, 첫 번째는 무작위 탐색 방법 두 번째는 DBAM과 ABAM의 융합 모델을 사용해 보았고, 세 번째는 6각형 기반의 Q-learning 알고리즘을 사용하였으며, 마지막으로는 본 논문에서 제안한 12각형 기반의 Q-learning과 Cascade SVM을 이용한 알고리즘을 사용하여서 시뮬레이션 하였다.

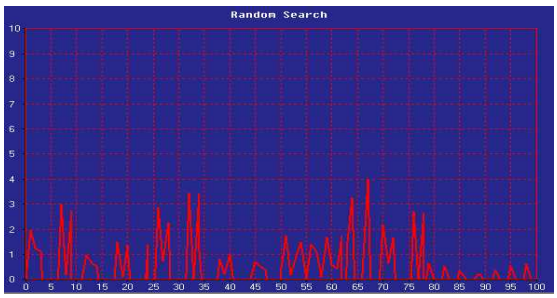

그림 8. 무작위 탐색
Fig. 8. Random search

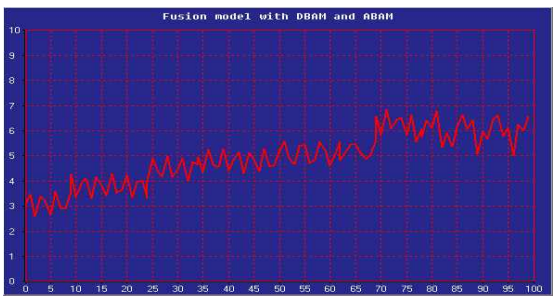

그림 9. DBAM과 ABAM의 융합모델
Fig. 9. Fusion model with DBAM and ABAM

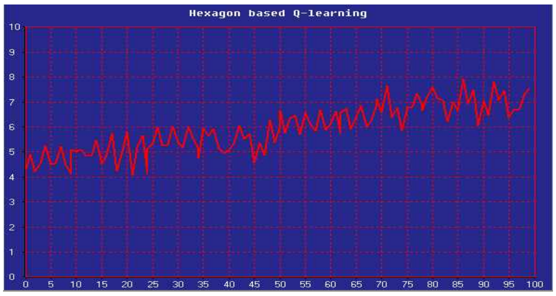

그림 10. 6각형 기반 Q-learning
Fig. 10. Hexagon based Q-learning

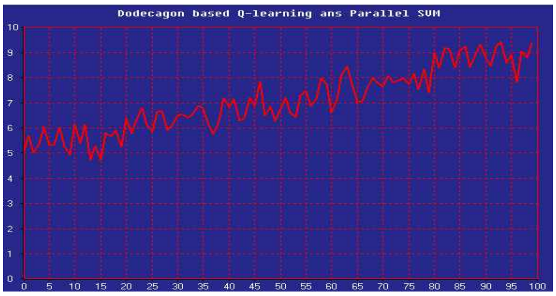

그림 11. 12각형 기반 Q-leaning과 Cascade SVM
Fig. 11. Dodecagon based Q-learning and Cascade SVM

총 100회의 탐색 시도에서 랜덤 탐색의 경우는 횟수가 늘어남에도 불구하고 특별한 규칙을 발견하기 힘들었다. 또한 랜덤 탐색의 특성상 통계적인 의미를 부여하기는 어려웠다. 다음으로 DBAM과 ABAM의 융합 모델인 경우, 모든 100회의 시행동안 평균적으로 4,5대 정도의 로봇이 목표물을 찾아내었다. 이것은 DBAM과 ABAM 융합모델을 통해서도 탐색의 성능이 상당히 강화 될 수 있음을 나타낸다. 세 번째 방법인 6각형 기반의 Q-learning 알고리즘은 평균적으로 6대 정도의 로봇이 목표물 탐색에 성공하였으며, 마지막으로 본 연구에서 제안한 12각형 기반 Q-learning과 Cascade SVM 알고리즘을 통한 탐색의 결과는 총 100회 시행에 평균적으로 7.5대 정도의 로봇이 목표물 탐색에 성공하였다. 본 논문에서 제안한 방법이 기존의 3가지 방법보다는 과거의 상황에 대해서 좀 더 정확성 있게 상황인식을 하는 결과를 가지고 왔다.

V. 결론

본 논문에서는 먼저 선험적 지식이 없고 장애물이 놓여있는 공간에서의 목표물 탐색 알고리즘으로 12각형 기반의 Q-learning과 Cascade SVM 알고리즘을 제안하였다. 제안한 알고리즘은 10대의 로봇과 25개의 장애물 그리고 1개의 목표물로 가정하여 실험을 하였고, 실험의 결과를 통해 12각형 기반의 Q-learning과 cascade SVM 알고리즘이 목표물 탐색에 있어서 새로운 방법이 될 수 있다는 것을 보여주었다.

향후 연구는 다음과 같은 점을 고려하여 진행해 나갈 예정이다. 첫째 다수의 로봇이 좀 더 효율적으로 목표물을 추적하기 위해서는 로봇간의 통신 기술이 뒷받침 해주어야 한다고 생각한다. 이에 따라 센서 네트워크를 이용한 통신 기법 혹은 Data Mining 기법에 관한 연구도 선행되어져야 할 것이다. 둘째 본 논문에서 제안한 로봇의 수 자체가 군집로봇 시스템 상의 가장 기본이 되는 개수이므로, 앞으로는 수 십대 혹은 수백 대의 로봇을 기본으로 시뮬레이션 해보는 작업도 진행해야 할 것이다. 마지막으로는 다수의 로봇이 실제 환경에서 하나의 목표물을 추적한다는 것은 시뮬레이션 상에서보다 더 많은 장애물과 방해요소가 존재하게 된다. 또한 장애물 자체도 정적으로 존재하는 장애물뿐만 아니라 동적으로 변하는 장애물에 대한 회피 알고리즘도 고려되어져야 할 것이다.

참 고 문 헌

[1] [illegible], "12각형 기반의 Q-learning과 SVM을 이용한 군집로봇의 목표물 추적 알고리즘," 한국지능시스템학회 논문지, 제 18권, 제 3호, pp. [illegible], 2008.

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

[3] G. Oggasawara, 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.

[4] D. Ballard, An Introduction to Natural Computation, The MIT Press Cambridge, 1997.

[5] Hans Peter Graf, Eric Cosatto, Leon Bottou, Igor Durdanovic, and Vladimir Vapnik, "Parallel Support Vector Machines : The cascade SVM," NIPS, 2004.

[16] Jian-Pei Zhand, Zhong-Wei Li, and Jing Yang, "A parallel SVM training algorithm on large-scale classification problems," Proceeding of the Fourth International Conference on Machine Learning and Cybernetics, pp. 1637-1641, 2005.
[17] [illegible], "SVM-SMO와 Pan-Tilt [illegible]," [illegible], 제31권, 제2호, vol. 31, no. 2, pp. 679-681, 200[illegible].
[18] W. Ashley, T. Balch, "Value-based observation with robot teams (VBORT) using probabilistic techniques," Proc. of Int. Conf. on Advanced Robotics, 2003.
[19] W. Ashley, T. Balch, "Value-based observation with robot teams (VBORT) for dynamic targets," Proc. of Int. Conf. on Intelligent Robots and Systems, 2003.
[20] Ying Wang, "Multi-robot box-pushing: single-agent Q-learning vs. Team Q-learning," Proc. of International Conference on Intelligent Robots and Systems, 2006.
[21] Guo-Sheng Yang, Er-Kui Chen, "Mobile robot navigation using neural Q-learning," Proceedings of the Third International Conference on Machine Learning and Cybernetics, 2004.
[22] Hyun-Chang Yang, Ho-Duck Kim, Han-Ul Yoon, In-Hun Jang, and Kwee-Bo Sim, "Hexagon-based Q-learning algorithm and applications," International Journal of Control, Automation, and Systems, vol. 5, no. 5, pp. 570-576, 2007.

저 자 소 개

서 상 욱

2007년 중앙대학교 전자전기공학부 공학사.
2007년~현재, 중앙대학교 전자전기공학부 석사과정.
관심분야 : Machine learning, Ggroup behavior and control of swarm robot, Multi agent robotic system 등.
Email : ssw0511@wm.cau.ac.kr

양 현 창

2002년 숭실대학교 산업공학과 공학석사.
2006년~현재, 중앙대학교 전자전기공학부 박사과정.
관심분야 : 군집로봇시스템, 유비쿼터스 센서 네트워크, 스마트 홈.
Email : icros@icros.or.kr

심 귀 보

1984년 중앙대학교 전자공학과 공학사.
1986년 중앙대학교 전자공학과 공학석사.
1990년 The University of Tokyo(Japan) 전자공학과 공학박사.
1991년~현재, 중앙대학교 전자전기공학부 교수.
관심분야 : 지능시스템, 유비쿼터스 지능형 로봇, 로봇의 감정인식 및 표현시스템, 지능형 감시시스템, 지능형 홈, 유비쿼터스 컴퓨팅 및 센서네트워크, 적응 및 학습알고리즘, 소프트 컴퓨팅, 다개체 및 자율분산로봇시스템, 진화하드웨어 및 임베디드시스템.
Email : kbsim@cau.ac.kr