양자 강화학습 기반 자율 수중 차량 제어

정재현, 박수현*, 김중헌
고려대학교, *숙명여자대학교

{rupang1234, joongheon}@korea.ac.kr, *soohyun.park@sookmyung.ac.kr

Quantum Reinforcement Learning-based Autonomous Underwater Vehicle Control

Jaehyun Chung, Soohyun Park*, Joongheon Kim
Korea Univ., *Sookmyung Women Univ.

요 약

양자 강화학습은 차원의 저주 등 기존 강화학습의 한계점을 보완한 강화학습이다. 본 논문은 자율 수중 차량의 어뢰 회피율을 높이기 위한 지능형 임무지원시스템을 연구하고, 양자 강화학습을 활용하여 효율적인 어뢰기만전술을 제안한다. 이는 성능평가를 통해 제안한 알고리즘의 성능이 효율적임을 나타낸다.

I. 서 론

최근 자율 수중 차량 시스템의 미래 전투 성능을 극대화하기 위해 지능형 임무지원시스템과 자동화 기술을 통합하는 연구가 활발히 이루어지고 있다. 보통 어뢰의 속도는 자율 수중 차량보다 훨씬 빠르기 때문에 긴급한 상황에서는 신속하고 정확한 계산이 필요하다. 따라서, 본 논문에서는 기존 강화학습에 양자 특성을 도입하여 어떤 상황에서도 대응 가능한 자율 수중 차량의 회피 침로를 선택할 수 있는 적 어뢰기만전술 알고리즘을 제안한다. 양자 강화학습(Quantum Reinforcement Learning, QRL)을 활용하면 학습 파라미터 수 감소를 통해 연산량이 감소하여 신속하게 대응할 수 있다. 본 논문의 성능평가를 통해 해당 알고리즘의 우수성을 평가한다.

II. 본론

2-1 어뢰기만전술

최근에는 자율 수중 차량이 기만기를 사용하여 어뢰를 회피하는 어뢰기만전술이 사용된다 [1]. 기만기는 자율 주행 차량이 회피기동을 할 수 없는 상황에서 소음을 발생하여 어뢰를 유도하는 장치이다. 이를 통해 자율 주행 차량의 생존율을 높인다. 기만기는 부유식 기만기(Hovering Decoy)와 자항식 기만기(Mobile Decoy)로 두 가지 종류가 있다 [2]. 그림 1 은 부유식 기만기와 자항식 기만기를 대략적으로 나타낸 그림이다.

부유식 기만기는 정지한 상태로 운용되는 기만기이다. 부유식 기만기는 수중에서 고정된 상태로 어뢰를 기만하는 역할을 수행하며 일정한 위치를 유지한다. 정지한 상태로 운용되는 특성 때문에 다수의 기만기를 동시에 운용하기도 한다.

자항식 기만기는 부유식 기만기와 달리 이동하는 특성을 지닌 기만기이다. 이는 어뢰가 접근하는 방향에 수직으로 위치한 경로에서 발사되도록 모델링되었다. 자항식 기만기는 잡음 신호를 발생시키고 어뢰의 위치를 감지하여 표적 신호를 모방하는 기능을 갖추고 있어, 이는 부유식 기만기의 한계를 보완하여 자함의 생존성을 더 높이게 되었다.

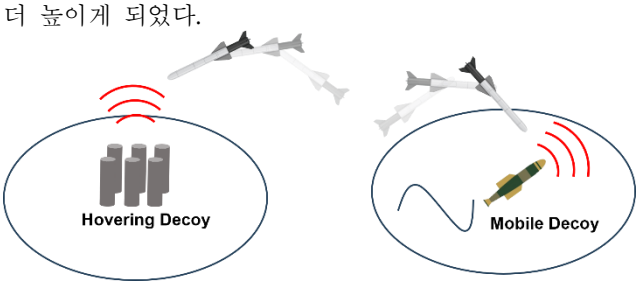

그림 1 부유식 기만기(왼) 및 자항식 기만기(오)

2-2 양자 강화학습

양자 컴퓨팅은 0 과 1 을 동시에 나타내는 큐비트를 사용한다 [3]. 그림 2 는 양자 인공신경망의 구조를 보여준다. 양자 인공신경망은 고전 데이터를 입력으로 받아 양자 상태로 인코딩을 진행하여, 매개변수화된 양자 회로에서 중첩과 얽힘을 생성하는 양자 게이트를 통해 큐비트의 상태를 변환하여 비선형성을 확보하여, 이러한 큐비트들의 상태는 측정을 통해 하나의 고전 데이터 값으로 결정한다 [4].

강화학습은 기계 학습의 한 분야로, 에이전트가 환경과 상호작용하며 보상을 최대화하기 위해 학습하는 과정을 다룬다 [5]. 이는 보상을 받는 행동을 선택하고 시행착오를 통해 학습하여 최적의 전략을 개발하는 것을 목표로 한다.

양자 강화학습은 기존 강화학습에 양자 인공신경망을 도입한 강화학습이다. 이를 통해 기존 강화학습이 해결하기 어려운 복잡한 문제를 양자 특성을 통해

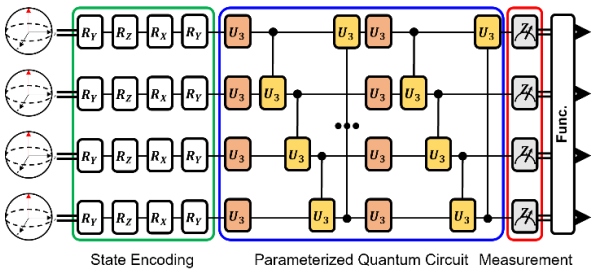

그림 2 양자 인공신경망

효율적으로 다룰 수 있다. 양자 강화학습은 기존 강화학습보다 필요한 학습 파라미터 수가 감소하여 학습 속도가 향상될 수 있으며, 양자 특성을 통한 행동 공간 차원 축소를 통해 복잡한 문제에서도 효율적인 솔루션이 제공이 가능해진다.

본 연구에서 상태는 자율 수중 차량 및 두 어뢰의 x, y 좌표로 6 가지이고, 행동은 자율 수중 차량의 이동 방향 및 기만기 발사 시 기만기의 이동방향으로 8 가지이다. 보상은 평소 자율 수중 차량과 두 어뢰 간 상대적 최소 거리이며 피격 시 -25 의 음의 값을 갖는 것으로 설계하였으며, 이에 따라 보상함수 R 은 다음과 같이 나타낼 수 있다.

$$R = \begin{cases} -25 & (if\ hit) \\ \dfrac{\min(d_1, d_2)}{d_0} & (otherwise) \end{cases}$$

여기서 d_1, d_2 는 각각 두 어뢰와 잠수함 사이 거리를 나타내며, d_0 는 초기 상황에서 두 어뢰와 잠수함 사이 최소 거리를 나타낸다.

2-3 실험 결과

본 장에서는 자율 수중 차량이 서로 다른 위치에 있는 두 어뢰를 회피하는 상황에서 실험을 진행하였다. 이때, 자율 수중 차량이 이동할 수 있는 범위는 상하좌우로 4 가지이며, 기만기 역시 마찬가지이다. 어뢰는 상하좌우 및 대각선 방향으로 8 가지이다. 그 이유는 자율 수중 차량이 피격되어야 할 상황에서 피격되지 못하는 경우가 발생할 수 있기 때문이다. 강화학습을 진행하기 위해서 상태, 행동, 보상을 정의해야 한다. QRL 의 대조군으로 Deep Q-Network (DQN)이 사용되었다. 실험은 총 5000 번의 Episode 로 진행하였다.

첫번째 지표는 알고리즘 성능이다. 그림 3 은 QRL 및 DQN 의 평균 보상 값을 나타내는 그래프이다. 두 알고리즘 모두 일정한 값으로 수렴하는 추세를 보인다. 이때, QRL 은 DQN 과 비교하였을 때 성능이 비슷하거나 조금 더 좋은 성능이 나왔다. 이처럼 성능 면에서 DQN 보다 QRL 이 우수하지만, 그 차이는 그렇게 크지 않는 것을 알 수 있다.

두 번째 지표는 모델 파라미터 수이다. 모델 파라미터 수는 각 알고리즘의 인공신경망에서 결정된다. 이때, DQN 에서는 입력 레이어 1 개, 출력 레이어 1 개, 히든 레이어 2 개를 사용하였으며, 각 계층에서 사용된 총 1632 개의 파라미터 수가 사용된다. 반면, QRL 에서는 6 개의 양자 게이트가 사용되며, 총 모델 파라미터 개수는 60 개이다. 이를 통해 QRL 이 DQN 보다 훨씬 적은 파라미터 수를 사용하며, 이는 QRL 이 DQN 보다 훨씬 효율적인 학습을 진행하는 것을 알 수 있다.

III. 결론

본 논문은 두 어뢰를 회피하는 간단한 시나리오에서 자율 수중 차량의 효율적인 어뢰기만전술 및 회피

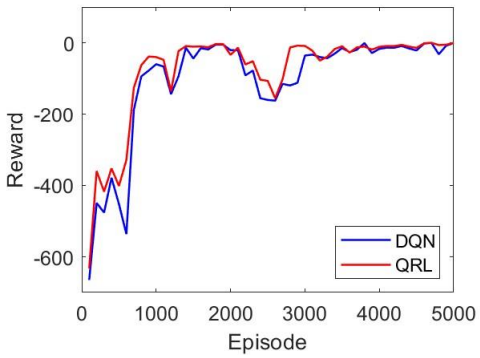

그림 3 평균 보상 그래프

침로를 위한 양자 강화학습 기반 자율 수중 차량 기동 방법을 제안한다. 본 연구는 어뢰와 자율 수중 차량이 이동할 수 있는 격자의 수를 다르게 하여, 어뢰가 자율 수중 차량 보다 빠르다는 사실을 함께 반영하였다. 본 논문에서는 해당 문제에 양자 강화학습과 기존 강화학습 비교를 통해 자율 수중 차량 회피기동 및 어뢰기만전술 성능을 평가하였다. 그 결과, 양자 강화학습이 기존보다 성능은 비슷하거나 더 높게 나왔으며, 파라미터 수가 더 적게 사용된 것을 알 수 있다. 이를 통해 더 복잡한 실제 환경에서도 자율 수중 차량이 효율적인 어뢰기만전술과 회피기동을 통해 생존율이 높아질 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 2022 년 한국연구재단의 지원을 받아 수행됨 (NRF 2022R1A2C2004869). 본 논문의 교신 저자는 김중헌임.

참 고 문 헌

[1] K. H. Liang, K. M. Wang, “Using Simulation and Evolutionary Algorithms to Evaluate the Design of Mix Strategies of Decoy and Jammers in Anti-Torpedo Tactics,” in *Proc. IEEE Winter Simulation Conference*, Monterey, CA, USA, December 2006, pp. 1299-1306.

[2] M. Shin, H. Cho, J. Lee, J. S. Lim, S. Lee, W. J. Kim, W. Hong, “Effectiveness Analysis for Survival Probability of a Surface Warship Considering Static and Mobile Decoys,” *Journal of the Korea Society for Simulation*, vol. 25, no. 3, pp. 53-63, May 2016.

[3] G. Brassard, I. Chuang, S. Lloyd, and C. Monroe, "Quantum Computing," *The National Academy of Sciences*, vol. 95, no. 19, pp. 11032-11033, September 1998.

[4] Y. Kwak, W. J. Yun, S. Jung, and J. Kim, “Quantum Neural Networks: Concepts, Applications and Challenges,” in *Proc. International Conference on Ubiquitous and Future Networks (ICUFN)*, Jeju Island, Korea, Republic of, August 2021, pp. 413-416.

[5] C. Park, G. S. Kim, S. Park, S. Jung, and J. Kim, “Multi-Agent Reinforcement Learning for Cooperative Air Transportation Services in City-Wide Autonomous Urban Air Mobility,” *IEEE Transactions on Intelligent Vehicles*, vol. 8 no. 8, pp. 4016-4030, June 2023.