
| | |
|--------------------|---|
| 저자 (Authors) | 박경근, 김현기 Kyoung Keun Park, Hyungi Kim |
| 출처 (Source) | 한국항공우주학회 학술발표회 초록집 , 2020.11, 1354-1355 (2 pages) |
| 발행처 (Publisher) | 한국항공우주학회 The Korean Society for Aeronautical & Space Sciences |
| URL | http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE10526658 |
| APA Style | 박경근, 김현기 (2020). 포텐셜 필드를 이용한 Deep Q-Learning 기반 무인기의 미사일 회피. 한국항공우주학회 학술발표회 초록집, 1354-1355. |
| 이용정보 (Accessed) | 한국외국어대학교 203.253.93.*** 2021/04/02 13:33 (KST) |

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

포텐셜 필드를 이용한 Deep Q-Learning 기반 무인기의 미사일 회피

박경근^{1*}, 김현기¹

한국항공우주산업¹

Missile Evasion of UAV based on Deep Q-Learning using Artificial Potential Field

Kyoung Keun Park^{1*}, Hyungi Kim¹

Key Words : Deep Q-Learning(딥큐러닝), Reinforcement Learning(강화학습), Potential Field(포텐셜 필드), Missile Evasion(미사일 회피)

서 론

6세대 전투기가 도입되는 미래 전장환경에서는 무인전투기가 전장에 활약할 가능성이 높다. 기존의 유인전투기가 수행하던 임무를 무인기가 수행하기 위해서는 다양한 무인기에 대한 운용 전략을 수립해야 한다. 무인 플랫폼의 운용 전략을 수립할 때 강화학습을 통해 컴퓨터가 상황을 학습하여 더 나은 전략으로 결정하는 방식에 대한 연구가 활발히 이루어지고 있다. 무인 전투기 분야에서도 강화학습을 기반으로 무인기의 1대1 교전 전략을 학습시키는 연구가 이루어졌다. ⁽¹⁾

공중전에서 미사일은 적을 타격하는 주요 수단이며 마찬가지로 아군 무인기가 미사일을 효과적으로 회피할 수 있어야 임무를 효율적으로 수행할 수 있다. 효율적인 미사일 회피 전략을 수립하기 위해 강화학습의 한 종류인 Q-Learning을 이용해 무인기 미사일 회피전략을 컴퓨터가 학습하는 방식에 대한 연구가 수행되었다. ⁽²⁾ 더 나아가 Deep Q-Learning을 이용해 미사일 회피전략을 더욱 효율적으로 학습시키는 방법에 대해서도 연구되었다. ⁽³⁾

효과적인 강화학습을 수행하기 위해서 학습 목적에 맞는 적합한 Reward를 적용해야 한다. 학습에 사용되는 Reward에 따라 학습 후 도출된 해의 성능이나 학습의 속도가 달라진다. 행동자(Agent)의 Potential 차이를 추가적으로 Reward에 반영함으로써 학습 성능을 향상시키는 연구가 진행되었다. ⁽⁴⁾

본 논문에서 무인기-미사일간 충돌 Potential Function을 정의하고 Potential 차이를 Reward로 사용하여 무인기가 미사일을 회피하는 전략을 Deep Q-Learning 기반으로 학습시키는 방식을 제시한다. 또한 간단한 시뮬레이션 환경을 구성하여 무인기의 미사일 회피 전략을 생성해 보았다.

이론적 배경

Agent가 State s 에서 Action a 를 수행하여 State s' 가 됐을 때의 Reward를 두 상태의 Potential 차이로 수식 1과 같이 사용한다. ⁽⁴⁾ 함수 P 는 Potential function 이며 γ 는 강화학습에서 사용되는 감쇠율(discount rate) 이다.

$$R(s, a, s') = \gamma P(s') - P(s) \quad (1)$$

항공기의 충돌회피를 위해 Potential Field를 정의한 연구가 있다. ⁽⁵⁾ 해당 연구에서 사용한 Potential function을 무인기의 미사일 회피 문제로 적용시킬 경우 수식 2와 3과 같다. R_{mis} 와 R_{UAV} 는 미사일과 무인기의 위치 R_{crash} 는 미사일이 무인기를 격추 시킬 수 있는 반경이다. R_{turn} 은 무인기가 회전할 수 있는 반경이며 θ 는 미사일과 무인기 간 상대거리와 상대속도간의 각도 차이이다.

$$P = \exp\left(-K \left(\frac{\|R_{mis} - R_{UAV}\|}{R_{crash}}\right)^2\right) \quad (2)$$

$$K = \left(1 - \left(\frac{R_{crash}}{R_{crash} + R_{turn}}\right)^2\right) \sin^2 \theta + \left(\frac{R_{crash}}{R_{crash} + R_{turn}}\right)^2 \quad (3)$$

수식 2의 Potential Function을 수식 4와 같이 변형해 사용했다. Reward의 정의에 따라 무인기와 미사일간의 거리가 멀수록 Potential이 커지도록 하고 미사일과 무인기간의 거리가 먼 경우에도 유의미한 Reward를 줄 수 있도록 변형했다.

$$P = -\left(K \left(\frac{\|R_{mis} - R_{UAV}\|}{R_{crash}}\right)^2\right)^{-1/2} \quad (4)$$

수식 1, 3, 4을 이용해 특정 State S 에서 행동 a 에 따른 Reward를 정의할 수 있다.

시뮬레이션

시뮬레이션 구성

무인기의 진행방향에서 미사일이 무인기를 격추시키기 위해 접근하는 2차원 시뮬레이션 환경을 구성하였다. 무인기와 미사일이 접근하는 초기 조건을 설정하고 무인기와 미사일이 운동학 법칙에 따라 이동하는 환경을 생성하였다. 제시한 Reward를 이용해 Deep Q-Learning을 통해 무인기의 회피기동을 학습시켰다.

무인기에게 접근하는 미사일은 수식 5와 같은 비례항법 (Proportional Navigation)을 사용한다고 가정하였고 가속에 제한을 두었다.

$$a_{mis} = NV_{mis}\dot{\lambda} \quad (5)$$

무인기는 정면으로 가속, 감속, 좌회전, 우회전, 직진의 다섯 가지 행동을 취할 수 있다고 가정하였다. 무인기 입장에서 관측되는 미사일의 State는 미사일까지의 거리, 상대 속도의 거리방향 성분, 각도, 각속도로 설정했다.

시뮬레이션 결과

Fig.1 은 첫 훈련 시나리오를 결과를 나타내며 파란색 선이 무인기의 궤적, 붉은색 선이 미사일의 궤적이다. 아직 학습이 되지 않아 무인기가 단조로운 회피기동을 했음을 알 수 있다.



Fig. 1. 학습 초기 무인기와 미사일의 궤적

Fig.2 는 Reward를 적용시켜 Deep Q-Learning을 통해 학습 시킨 후 학습된 회피기동을 수행한 결과이다. 제시한 학습 방식을 통해 컴퓨터가 회피 전략을 생성해 회피기동을 수행했음을 알 수 있다.



Fig. 2. 학습 후 무인기와 미사일의 궤적

결론

본 논문에서는 무인기가 미사일을 회피하는 전략을 Deep Q-Learning 기반으로 학습 할 때 Potential Function의 차이를 Reward로 사용하는 방식을 제시하였다. 제시한 방법으로 학습할 경우 미사일에 대한 무인기의 회피 궤적을 생성할 수 있음을 컴퓨터 시뮬레이션을 통해 보였다.

추후 연구에서는 전장 환경을 3차원으로 발전시키고 무인기와 미사일의 센서, 기동 등을 실제와 가깝게 모델링 하여 실제 무인기에 탑재 할 수 있는 알고리즘으로 발전시킬 것이다. 또한 여러 학습 방법을 통해 도출된 회피 성능을 비교하여 회피를 가장 잘 수행할 수 있는 알고리즘을 찾을 것이다.

참고문헌

- 1) Yesheng Zhang, Wei Zu, Yang Gao, Hongxing Chang, "Research on Autonomous Maneuvering Decision of UCAV Based on Deep Reinforcement Learning," *2018 Chinese Control And Decision Conference*, 2018, pp.230-235.
- 2) Dongjin Lee, Hyochoong Bang, "Missile Evasive Strategies for Unmanned Aircrafts using Reinforcement Learning," *Proceeding of the 2012 KSAS Spring Conference*, 2012, pp. 470-475.
- 3) Jae-Eun Ryu, Kyujin Moon, Chang-Kyung Ryoo, "UAV Guidance to Evade Guided Missiles Based on Deep Q-Learning," *Proceeding of the 2018 KSAS Spring Conference*, 2018, pp. 225-226.
- 4) Andrew Y. Ng, Daishi Harada, Stuart Russell, "Policy invariance under reward transformations: Theory and application to reward shaping," *International Conference on Machine Learning*, 1999, pp. 278-287.
- 5) Thomas J. Stastny, Gonzalo A. Garcia, Shawn S. Keshmiri, "Collision and Obstacle Avoidance in Unmanned Aerial Systems Using Morphing Potential Field Navigation and Nonlinear Model Predictive Control," *Journal of Dynamic Systems Measurement and Control*, Vol. 137(1), 2015.