

드론의 PID 제어와 강화학습 제어에 관한 연구 동향

Research Trends on PID Control and Reinforcement Learning Control on Drone

김 주 봉(Jubong Kim)*, 황 규 영(KyoYong Hwang)*,
안 채 현(Chae-Hun An)**, 한 연 희(Youn-Hee Han)*

*Department of Computer Science and Engineering
Korea University of Technology and Education
{hjkim, shyeo, yhhan}@koreatech.ac.kr

** Department of Mechatronics Engineering
Korea University of Technology and Education
ach@koreatech.ac.kr

요 약

본 논문은 다양한 드론 중에서 쿼드콥터 제어에 대한 롤링, 피칭, 요잉을 기본으로 하는 기본적 동역학 이론을 제시하고, 쿼드콥터의 전통적인 PID 제어와 강화학습을 이용한 제어에 관한 최신 연구 경향을 설명한다. 특히, 참고 문헌 [12]의 연구를 토대로 강화학습 아키텍처 및 쿼드콥터 제어를 위한 상태, 행동, 보상의 정의에 대해 설명한다. 마지막으로 PID 제어와 강화학습 제어의 비교 분석을 제공함으로써 향후 드론의 다양한 지능적 제어를 위하여 보다 심도있는 강화학습 제어 연구의 필요성을 제시한다.

키워드: 무인항공기, 드론, 군집드론, PID 제어, 인공지능, 강화학습

Abstract

This paper presents the basic dynamics theory based on rolling, pitching and yawing of quadcopter control, and explains the latest research trends on the conventional control using quadcopter's PID control and reinforcement learning control. In particular, we describe the definition of states, actions, and rewards for reinforcement learning architecture for quadcopter control based on the work of [12]. Finally, by providing a comparative analysis of PID control and reinforcement learning control, we propose the necessity of deeper reinforcement learning control study for various intelligent control of drones in the future.

Key words: Unmanned Air Vehicle, Drone, Drone Swarm, PID Control, Artificial Intelligence, Reinforcement Learning

1. 본 연구는 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 결과임 (No. 2018R1A6A1A03025526)
2. 교신저자: 한연희

1. 서론

드론이란 조종사의 탑승 없이 무선 전파의 유도 및 자체 제어기에 의해 비행하는 비행기나 헬리콥터 모양의 비행체인 ‘무인항공기’를 일컫는다[1]. ‘드론(drone)’이라는 단어의 원래 의미는 꿀벌의 수컷을 뜻하며, 작은 무인항공기의 날아다니는 모습이 꿀벌과 유사하여 드론이라는 이름이 붙여졌다.

최근 드론은 학계, 군, 각종 산업분야에서 매우 주목하고 있으며, BI Intelligence의 조사 보고서에 따르면, 2021년까지 드론 판매는 120억 달러를 넘을 것이며, 산업 드론 시장의 연평균 성장률은 51%에 달할 것이라고 한다[2]. 또한, 미국 방산업 전문컨설팅 업체인 틸그룹은 세계 드론 시장 규모가 2024년 147억 달러 규모로 급증할 것으로 예상하고 소형 드론 시장은 급성장해 전체 시장 규모가 2024년까지 연평균 15% 성장할 것이라고 전망하고 있다[3]. 그 중의 대부분이 군사용으로 많이 사용되고 있지만 각국의 법이 확립된다면 세계적으로 시장이 빠르게 성장할 것으로 전망된다.

드론은 날개의 형태에 따라 고정익기, 회전익기 그리고 두 가지 방식이 혼합된 복합형으로 나뉜다. 고정익기 비행체의 날개는 고정되어 있으며, 일반적인 비행기의 비행 양력 이론을 바탕으로 기체를 하늘로 띄운다. 흔히 ‘헬리콥터’로 알려져 있는 회전익기는 프로펠러라고도 불리는 로터(rotor)가 회전하면서 양력을 발생시키며 크게 단일로터(헬리콥터), 멀티로터(콥터)로 구분할 수 있다. 또한, 프로펠러의 숫자에 따라 바이콥터(2개), 트리콥터(3개), 쿼드콥터(4개), 헥사콥터(6개), 옥토콥터(8개)로 분류한다. 이 중 근래에는 쿼드콥터가 가장 주목받고 있으며, 본 논문에서는 주로 쿼드콥터의 제어 이론에 대해서 설명한다.

쿼드콥터는 기체 구조가 대칭적이어서 각 프로펠러의 속도제어만으로 충분히 방향 전환이 가능해 헬리콥터보다 구조적으로 훨씬 단순하다. 또한 기체 자체를 제어하기가 쉬워 누구나 날리기 편하고 다양한 용도로 사용이 가능하다. 쿼드콥터가 최근에 학

계, 군, 각종 산업 분야에서 사용이 가능해진 것은 2000년 이후에 쿼드콥터에 사용되는 센서, 모터 등의 부품이나 전자소자들이 소형화되고 가격이 낮아졌기 때문이다.

최근, 드론의 하드웨어적 기술이 비교적 안정화 되고 있는 상황에서, 소프트웨어적으로 드론을 지능적으로 제어하는 연구가 계속되고 있다. 본 논문은 쿼드콥터 제어에 대한 기본적 동역학 이론부터 시작하여, 전통적인 PID 제어와 강화학습을 이용한 쿼드콥터 제어에 관한 최신 연구 경향을 정리하면서 PID 제어와 강화학습 제어의 비교 분석을 제시한다.

2. 쿼드콥터 드론의 비행 동역학

드론은 2가지의 기본 힘인 양력과 회전력을 사용하여 비행한다. 헬리콥터의 경우 하나의 메인로터(Main rotor)가 동력 장치의 회전력을 받아 회전하면서 양력을 발생시켜 상승한다. 뉴턴의 운동 제3법칙 ‘작용과 반작용의 원리’가 작용하여 메인 로터가 회전하면 기체는 메인 로터가 회전하는 반대 방향으로 돌아가는 힘을 받는다. 따라서 단일로터 기체는 기체의 회전을 막기 위해 테일로터(Tail rotor)를 기체의 꼬리 부분에 만들어, 메인로터에 의해 생기는 기체의 회전력을 상쇄시킨다.

비행체는 일반적으로 가로, 세로, 수직의 3가지 운동축을 가진다. 이 세 축을 중심으로 비행체는 롤링(rolling), 피칭(pitching), 요잉(yawing)의 운동을 한다. [그림 1]에서 볼 수 있듯이, 비행체의 동체가 비행방향(Forward-Backward) 축을 중심으로 좌우(Right-Left)방향으로 회전하는 것을 롤링(횡전), 좌

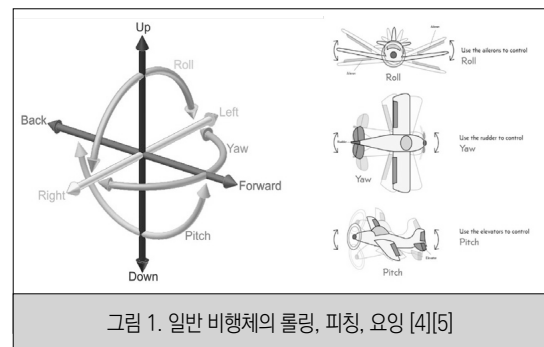


그림 1. 일반 비행체의 롤링, 피칭, 요잉 [4][5]

우축을 중심으로 회전하여 기수를 올리거나 내리는 것을 피칭, 수직(Up-Down)축을 중심으로 선회를 위해 좌우 지향점을 변경하는 것을 요잉이라고 한다.

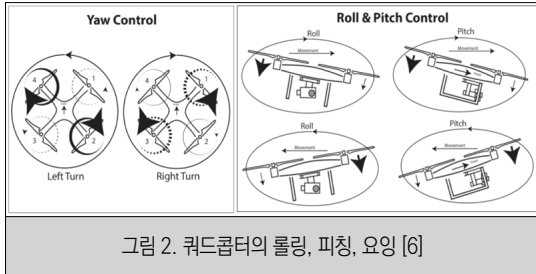


그림 2. 쿼드콥터의 롤링, 피칭, 요잉 [6]

쿼드콥터는 4개의 로터에서 나오는 회전력 w_0, w_1, w_2, w_3 을 이용하여 3개의 병진(translation) 운동과 3개의 회전(rotation)운동의 6자유도 (degree-of-freedom)로 움직이는 시스템이다. 각 3개의 회전은 롤링, 피치, 요잉이라고 불리우며, 4개의 로터의 회전력은 3개의 오일러 각도 ϕ, θ, ψ 및 상승추력(upward thrust) 정도인 f 에도 영향을 미친다.

[그림 2]에서 볼 수 있듯이 쿼드콥터의 프로펠러 1과 3은 같은 방향으로 회전하지만 2와 4는 반대 방향으로 회전한다. 쿼드콥터의 상승과 하강은 4개의 모든 프로펠러를 모두 같은 속도로 빠르게 회전시키거나 느리게 회전시키면 된다. 이 때 두 개의 마주보는 프로펠러 쌍씩(위 그림에서 프로펠러 1과 3, 그리고 프로펠러 2와 4) 서로 다른 방향으로 회전하기 때문에 반작용의 균형이 맞추어지면서 헬리콥터 처럼 불필요한 회전을 막기 위한 꼬리날개 같은 장치는 필요없다. 한편, 프로펠러 2, 4의 회전속도가 프로펠러 1, 3의 회전속도보다 빠르면 반작용의 균형이 무너지면서 쿼드콥터를 반시계방향으로 회전시킨다 (Left Turn). 반대로 프로펠러 1, 3의 회전속도가 프로펠러 2, 4의 회전속도보다 빨라지면 쿼드콥터를 시계방향으로 회전시킨다 (Right Turn). 쿼드콥터를 전진 및 후진하도록 하는 제어를 피치 제어라고 하며, 뒤쪽에 위치한 프로펠러 2개의 회전속도를 증가시키면 뒷부분이 앞부분에 비해 상대적으로 위로 들리면서 쿼드콥터를 앞으로 기울게 하

고, 이 자세의 기울기에 의해 앞으로 진행되는 추진력이 발생되어 쿼드콥터가 전진하게 된다. 후진이나 좌·우로의 이동도 이와 동일한 방법으로 제어된다. 지금까지의 설명을 기반으로 4개의 로터 속력과 4개의 운동 영향간의 관계를 수식으로 나타내면 다음 식 (1)과 같다.

$$\begin{aligned}
 u_f &= b(w_1^2 + w_2^2 + w_3^2 + w_4^2) \\
 u_\phi &= b(w_1^2 + w_2^2 - w_3^2 - w_4^2) \\
 u_\theta &= b(w_1^2 - w_2^2 - w_3^2 + w_4^2) \\
 u_\psi &= b(w_1^2 - w_2^2 + w_3^2 - w_4^2)
 \end{aligned}
 \tag{1}$$

[그림 3]은 지금까지 설명한 쿼드콥터의 운동을 외부에서 명령을 인가할 수 있는 수동 제어기 예시이다.

운동 이름	설명
Throttle	드론의 수직 상승 및 하강 제어
Yaw	드론의 시계 방향 및 반시계 방향 회전 제어
Pitch	드론의 전진 및 후진 이동 제어
Roll	드론의 좌향 및 우향 이동 제어

그림 3. 쿼드콥터의 수동 제어기 [6]

3. 쿼드콥터 드론의 PID 제어

쿼드콥터는 자이로, 가속도 등의 센서 값을 기반으로 조종기 명령의 입력으로 로터를 제어할 때 동시에 피드백 루프 (Feedback-Loop) 제어를 한다. 수동 제어기와 같은 외부 제어기에서 쿼드콥터에 명령을 주었을 때 쿼드콥터가 안정적이고 정밀하게 비행되는 이유는 자이로, 가속도 등의 센서와 마이크로 컨트롤러 보드에서 수행되는 피드백 루프 제어 방식

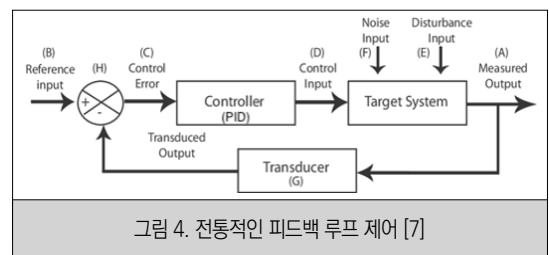


그림 4. 전통적인 피드백 루프 제어 [7]

에 기인한다.

[그림 4]에서 볼 수 있듯이 피드백 루프 제어는 대상 시스템(Target System) 동작의 측정 출력(A. Measured Output)과 참조 입력(B. Reference Input)을 비교하여 제어 오류(C. Control Error)를 산출한 다음 대상 시스템이 제어 시스템에 의해 정의된 대로 동작하도록 제어 입력(D. Control Input)을 계산한다. 측정 출력은 외부의 외란 입력(E. Disturbance Input) 또는 시스템 적응 자체에 의해 야기되는 노이즈 입력(F. Noise Input)에 의해 영향을 받을 수 있다. 변환기(G. Transducer)는 오류 계산에 용이하도록 센서에서 신호값을 변환한다[8].

예를 들어, 쿼드콥터의 고도(Measured Output)가 원하는 고도(Reference Input)보다 낮으면 설정된 고도값과 실제 고도값의 차이(Control Error)는 플러스가 될 것이고 로터 회전 속도를 증가(Control Input)시켜서 쿼드콥터를 좀 더 상승시킨다.

이와 같은 피드백 루프 제어의 중앙에 위치한 제어기(Controller)에 사용되는 제어 방식으로서 가장 흔히 사용되는 것은 PID 제어이며, 아래 식 (2)과 같이 시각 t 에서의 비례(Proportional) 제어인 P제어, 적분(Integral) 제어인 I제어, 미분(Derivative)제어인 D제어에 의해 조정되는 값(Manipulated Value, MV)들의 합으로 구성된다.

$$MV(t) = P(t) + I(t) + D(t) \\ = K_p e(t) + K_i \int_{t'=0}^t e(t') dt' + K_d \frac{de(t)}{dt} \quad (2)$$

위 식 (1)에서 $e(t)$ 는 시각 t 에서의 오차이며, K_p , K_i 및 K_d 는 제어 파라미터이며 이득값 또는 게인(Gain)이라고 하고, 적절한 이득 값을 수학적 혹은 실험적/경험적 방법을 통해 계산하는 과정을 튜닝(tuning)이라고 한다. PID 제어기의 튜닝에는 여러 가지 방법들이 있는데, 그중 가장 널리 알려진 것으로는 지글러-니콜스(ziegler-Nichols) 방법이 있다[9]. P 제어(비례제어)는 오차에 비례한 조정을 하는 기본적인 제어이다. 하지만 비례제어 만으로는 정상상태(steady-state)에서의 오차를 줄일 수 없으며, I 제어(적분 제어)가 과거 오류 변화를 고려하여 그러한 오

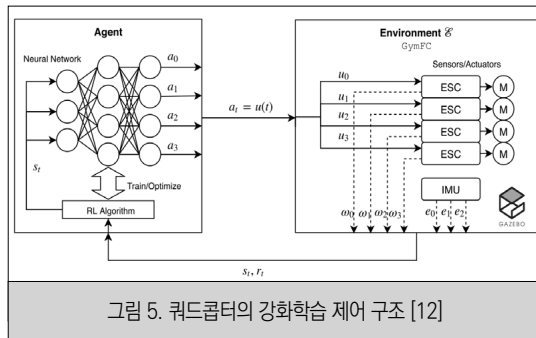
차를 줄여준다. D 제어(미분 제어)는 미래 오류 변화를 고려하여 출력값의 급격한 변화에 제동을 걸어 오버슈트(overshoot)을 줄이고 안정성(stability)을 향상시킨다[10][11].

쿼드콥터에서 일반적으로 P 제어는 조정기 스틱에 얼마나 빠르게 반응하는지 결정한다. P제어에 의한 조정값이 너무 높으면 갑자기 밝은 페달처럼 목표 속도에 빠르게 도달하겠지만 쉽게 그 목표 속도를 넘겨버릴 수도 있다. 이에 반응하여 어떤 방법으로도 다시 속도를 급히 낮추면 다시 곧바로 속도를 높여야 하기 때문에 드론이 급하게 떨리게 된다. 반면 P제어 조정값이 너무 낮으면 조정기 스틱에 대한 반응이 너무 느려서 드론이 흐느적거릴 수 있다. I제어는 드론의 회전이 멈출 때 원하는 회전보다 더 회전하거나 덜 회전하는 현상과 관계가 있다. 같은 드론이라도 배터리가 부착된 위치나 추가적으로 부착한 센서의 무게에 따라서도 I제어 조정이 필요하다. D제어는 목표한 속도에 드론이 얼마나 매끄럽게 반응하는지 조절한다. 갑자기 회전을 시작하거나 정지하면 드론의 행동이 경직되게 되며, D제어에 의한 조정값이 높을수록 빠르게 반응하는 드론이 되지만 그만큼 로터가 과열된다. D제어에 의한 조정값이 변하면 적절한 P제어 조정값이 함께 달라지기 때문에 보통 P제어 조정값을 먼저 정한 다음 D 제어 조정값을 조정한다.

쿼드콥터의 제어에는 롤링, 피치 및 요잉의 각 회전마다 PID 제어가 있다. 각 시각 t 에서 PID 합계가 각 축에 대하여 계산된 후, 이 값들은 Mixing이라는 프로세스를 통해 각 로터에 전달되는 전력량으로 변환된다. 각 로터 i 에 전달되는 제어 신호를 y_i 라고 정의할 때 y_i 는 다음 식 (3)과 같이 정의된다.

$$y_i = f(MV_{i,\phi} u_\phi + MV_{i,\theta} u_\theta + MV_{i,\psi} u_\psi) \quad (3)$$

4. 쿼드콥터 드론의 강화학습 제어



본 논문에서는 [12]에서 제안한 신경망 기반 컨트롤러로 구성된 [그림 5]의 강화학습 아키텍처를 중심으로 쿼드콥터의 강화학습 제어를 설명한다. 참고 문헌 [12]에서는 Gazebo 시뮬레이터[13] 기반의 고충실도(High-fidelity) 물리 시뮬레이션 환경 E를 구성하고 Irish quadcopter[14]와 상호 작용하는 강화학습 에이전트를 소개하였다.

각각의 스텝 t에서, 에이전트는 환경 E로부터 관찰(observation) o_t 를 수신한다. 관찰 o_t 는 쿼드콥터의 IMU(Inertial Measurement Unit)로부터 얻어지는 각 3개의 축에서의 각속도 오차 $\{e_0, e_1, e_2\} = \Omega^* - \Omega$ 와 ESC(Electric Speed Controller) 센서로부터 얻어지는 4개 로터의 각속도 $\{w_0, w_1, w_2, w_3\}$ 로 구성된다. 따라서, 각 관찰 o_t 는 연속 공간 $R^{7(=4+3)}$ 에 존재하는 임의의 벡터가 된다. 이러한 오차 식에서 Ω^* 는 3개의 축에 대한 제어의 최종 타겟 각속도 벡터이며 Ω 는 현재 각속도 벡터이다.

에이전트가 관찰 벡터 o_t 를 받으면 연속 공간 R^4 에 존재하는 벡터인 행동(action) a_t 를 도출한다. 이러한 행동 벡터는 ESC를 통해 각 4개의 로터에 전달되는 제어 신호 $u(t)$ 에 대응된다. 에이전트가 행동을 수행하면 환경 E로부터 해당 행동의 품질 및 성능을 나타내는 보상(reward) r_t 를 받는다. 정확히, 각 스텝 t마다 에이전트가 환경으로 받는 정보는 1) 보상 r_t 및 2) 과거 i개의 관찰 및 행동 정보를 포함한 상태 s_t 이다. 즉, 상태 s_t 는 다음과 같이 정의된다.

$$s_t = \{o_{t-1}, a_{t-1}, \dots, o_{t-1}, a_{t-1}, o_t\}$$

에이전트와 E 간의 상호 작용은 마코프 결정 프로세스(Markov decision process, MDP)로 정의된다. MDP 정의에서 주어진 현재 스텝 t의 상태와 행동이 각각 s_t 와 a_t 일 때 다음 스텝 t+1에 상태가 s' 로 천이되는 확률은 $\Pr\{s_{t+1}=s' | s_t=s, a_t=a\}$ 이다. 에이전트의 행동은 정책 π 에 의해 정의되며, 이는 본질적으로 특정 상태에 대해 취해야 할 행동을 매핑하는 것이다. 에이전트의 목적은 전체 스텝동안 받게되는 누적 보상이 최대가 되도록 최적의 정책 π^* 을 개발하는 것이다.

강화 학습 설계 시에 보상 r_t 의 정의가 가장 어렵다. 목적으로 하는 제어 포인트로 제어가 잘 되고 있는지를 정량적으로 정확하게 나타내도록 보상함수가 정의되어야 한다. [12]에서는 스텝 t에서의 보상 함수를 다음 식 (4)로 정의하고 있다.

$$r_t = -clip\left(\frac{sum(\Omega^* - \Omega)}{3V_{max}}, 0, 1\right) \quad (4)$$

위 식 (4)에서 V_{max} 는 3개의 축이 지닐 수 있는 최대 각속도를 의미하며, clip은 보상 값이 너무 작거나 클 경우 0과 1로 사이로 한정하는 함수이다. 식 (4)에 따르면 보상 값은 항상 음수이며, 강화학습의 목적은 누적 보상을 최대화하는 것이므로 에이전트는 가능한 빠른 시간 내에 오차를 줄이도록 학습이 될 것이다.

또한, 연속 행동 벡터를 산출하는 최적의 정책 개발 RL 기법 개발은 쉽지 않다. 연속 행동 벡터 산출이 가능한 현재까지 알려진 가장 성능이 좋은 RL 알고리즘은 PPO (Proximal Policy Optimization)이다. PPO 강화학습의 자세한 사항은 참고 문헌 [15]를 참고하도록 하자.

5. PID 제어와 강화학습 제어의 비교

PID 제어 방식은 지금까지 드론 레이싱을 포함하여 드론의 여러가지 행동 제어에 매우 많이 활용되고 있으며, PID 제어의 정확성(Accuracy) 및 안정성(Stability)은 오랜 기간 동안 검증되어져 왔다. 하

지만, 예기치 못한 환경, 즉 바람이 갑자기 분다거나 배터리가 갑자기 소진되어지는 상황에 대한 대처 능력은 그리 좋지 못하다. 즉, 시시각각 변하는 환경 및 동역학적 요소에 PID 제어는 그리 좋은 대안이 될 수 없다. 게다가, PID 제어 모델에 사용되는 3가지 게인(Gain) 값 조정을 튜닝이라고 하는데, 최적의 게인 값을 산출하는 튜닝 작업은 무척 어렵다. 최적의 게인 값을 산출하기 위한 좀 더 지능적인 튜닝 방법들이 많이 제안되고 있지만 실제적으로는 여전히 사람이 직접 몇몇 도구를 사용하여 휴리스틱하게 얻어낸다. 따라서, 이러한 PID 제어의 단점을 극복할 수 있는 대안으로 강화학습이 사용될 수 있다. PID 제어와 비교하여 강화학습이 지닌 장점을 요약하면 다음과 같다 [16].

- 드론의 동역학적 요소에 대한 어떠한 가정 및 모델없이 최적의 제어 정책을 만들 수 있다.
- 다양한 드론 형태와 센서, 로터 조합에 대해서 PID 제어는 사전에 수학적 모델 및 가정을 충분히 고려하며 설계되어야 하지만, 강화학습은 일관된 방식으로 상태, 행동, 보상에 대한 정의만 올바르게 이루어진다면 최적의 제어 정책 산출이 비교적 쉽다.
- 시시각각 변하는 예기치 못한 환경에 대한 대처 능력이 PID 제어에 비해 상대적으로 우수하다.

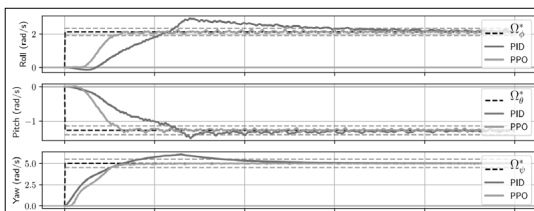


그림 6. PID 제어와 PPO 제어의 성능 비교 [12]

특히, 전통적인 PID 제어와 강화학습의 대표적인 알고리즘은 PPO에 대한 성능 분석 결과로서 참고 문헌 [12]에서 제시한 [그림 6]은 많은 시사점을 보여준다. 이 그림에서 볼 수 있듯이 쿼드콥터의 3개의 축에 대한 제어의 최종 타깃 각속도인 Ω^* 벡터 값

을 PID 제어 보다 PPO가 좀 더 빠른 시간 안에 달성하면서도 안정적으로 그 값을 유지하는 것을 볼 수 있다.

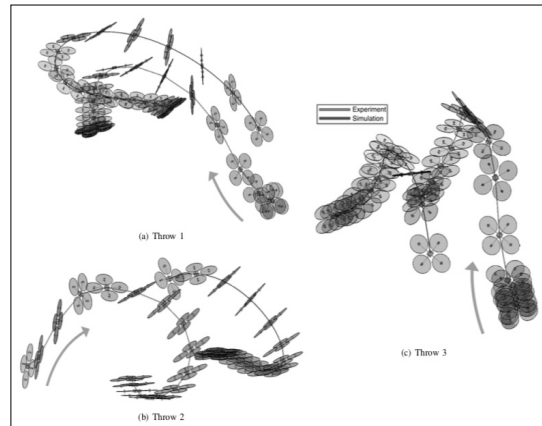


그림 7. 시뮬레이터와 실제계에서 강화학습 모델이 내장된 드론 던지기 이후의 이동 궤적 [16]

마지막으로 강화학습은 많은 횟수의 학습 과정이 필요하므로 시뮬레이터 또는 디지털트윈을 활용한 학습 이후 학습이 어느정도 완료된 모델을 다시 실제계의 드론에 내장된 컴퓨터 보드에 옮겨서 약간의 튜닝과 함께 테스트하는 작업이 필요하다. 참고 문헌 [16]에서는 드론을 던진 이후에도 자세를 올바르게 잡는 실험을 위하여 먼저 Robotic Artificial Intelligence (RAI) 시뮬레이터 [17]에서 Policy 최적화 강화학습을 진행한 후 학습이 완료된 모델을 Ascending Technologies에서 제작한 Hummingbird Quadrotor [18]에 탑재하여 드론의 운동 궤적을 [그림 7]과 같이 제시하였다.

6. 결론

본 논문은 여러 드론 중에서도 쿼드콥터 제어에 대한 롤링, 피칭, 요잉을 기본으로 하는 기본적 동역학 이론부터 시작하여, 전통적인 PID 제어와 강화학습을 이용한 쿼드콥터 제어에 관한 최신 연구 경향을 정리하였다. 특히, 참고 문헌 [12]의 연구를 토대로 강화학습 아키텍처 및 쿼드콥터 제어를 위한 상태, 행동, 보상의 정의에 대해 다루었다. 마지막으로 PID 제어와 강화학습 제어의 비교 분석을 제시하였

다. 앞으로 드론의 상업 및 군사적 이용 가치가 높아지면서 단순한 자세 및 운동 제어에서부터 복잡한 미션 제어까지 제어의 복잡도가 지속적으로 높아질 것으로 예상된다. 강화학습은 내재적으로 애기치 못한 상황에 대한 대처 능력이 PID 제어 보다 우수하므로 다양한 드론에 대한 보다 지능적인 제어를 위한 강화학습 연구가 더욱 지속되어야 할 것이다.

참고 문헌

- [1] 진정희, 이귀봉, “무인기/드론의 이해와 동향,” 한국통신학회지(정보와통신), 제33권 제2호, 80 - 85 (6 pages), 2016.01
- [2] BI Intelligence. The Drones Report: Market Forecasts, key players and use cases, and regulatory barriers to the proliferation of drones, 2016
- [3] S. Chae and S. Hong, “The Opportunity and Risk Analysis of Commercial Drones,” IITP, Jeju, Korea, June, pp. 736-737, 2015.
- [4] What is Pitch, Roll and Yaw?, <https://emissarydrones.com/what-is-roll-pitch-and-yaw>
- [5] Roll, Pitch, and Yaw, <https://howthingsfly.si.edu/flight-dynamics/roll-pitch-and-yaw>
- [6] How to Fly a Drone - The Ultimate Guide, <http://fromwhereidrone.com/how-to-fly-a-drone-the-ultimate-guide/>
- [7] N. Villegas, et al, “DYNAMICO: A Reference Model for Governing Control Objectives and Context Relevance in Self-Adaptive Software Systems,” Springer LNCS 7475, Aug. 2012.
- [8] H. Muller, M. Pezze, M. Shaw, “Visibility of Control in Adaptive Systems,” 2nd International Workshop on Ultra-Large-Scale Software-Intensive Systems (ULSSIS2008), 23-26, 2008
- [9] J. G. Ziegler, N. B. Nichols, “Optimum settings for automatic controllers” Transactions of the ASME, 64, pp. 759-768, 1942.
- [10] L. Uk, Y. Hun, J. Wook, P. Bae, and C. Ho, "Modeling and Controller Design of Quadrotor UAV," Proc. of the Korean Institute of Electrical Engineers Summer Annual Conf., Youngpyoung, Korea, July 2011, pp. 1922-1923.
- [11] B. Yun, J. Jeong, S. Ha, K. Kim, and J. Park, "Quad Copter attitude control with gyro sensor and acceleration sensor," J. of The Korean Society of Computer and Information, vol. 21, no. 1, 2013, pp. 111-112.
- [12] W. Koch et. al, “Reinforcement Learning for UAV Attitude Control,” ACM Transactions on Cyber-Physical Systems archive, Vol. 3, Issue 2., March 2019.
- [13] N. Koenig and A. Howard, “Design and use paradigms for Gazebo, an open-source multi-robot simulator,” Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'04), pp.2149-2154, Los Alamitos, CA, 2014.
- [14] APM Copter. 2018. Iris Quadcopter. Retrieved January 20, 2019 from <http://www.arducopter.co.uk/iris-quadcopter-uav.html>.
- [15] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” arXiv:1707.06347, 2017.
- [16] J. Hwangbo, I. Sa, R. Siegwart, and M.

Hutter, "Control of a Quadrotor with Reinforcement Learning," IEEE Robotics and Automation Letters 2, 4 pp. 2096-2103, 2014.

- [17] <https://bitbucket.org/leggedrobotics/rai>
- [18] M. Achtelik, M. Achtelik, S. Weiss, and R. Siegwart, "Onboard IMU and Monocular Vision Based Control for MAVS in Unknown In-and Outdoor Environments," in Robotics and Automation (ICRA), 2011 IEEE international conference on. IEEE, 2011, pp. 3056-3063.