

# 도심 항공 택시 비행 안정성 확보를 위한 강화학습 기법 연구

김준영\*, 신민규\*, 조준형, 정소이

아주대학교 AI 융합네트워크학과\*, 아주대학교 전자공학과

{\*junzero0615, \*saycode99, wngudd0126, sjung}@ajou.ac.kr

## Enhancing Urban Air Mobility Flight Stability based Reinforcement Learning Approach

Junyoung Kim\*, Mingyu Shin\*, Junhyung Cho, Soyi Jung

\*Dept. of Artificial Intelligence Convergence Network, Ajou University,  
Dept. of Electrical and Computer Engineering, Ajou University

### 요약

차세대 교통수단으로 도심 항공 교통(urban air mobility, UAM) 운용이 검토되고 있다. UAM은 기체 크기가 상용화된 드론에 비해 크기가 매우 커 기존 드론의 제어 방식으로는 동적으로 변화하는 환경에서 UAM 기체 자세 제어를 효과적으로 수행하기 어렵다. 이에 본 논문에서는 UAM의 안정적인 기체 자세 제어를 유지하기 위하여 deep deterministic policy gradient(DDPG) 강화학습 기법에 기반해 새로운 제어 방식을 제안한다.

### I. 서론

최근, unmanned aerial vehicles(UAV)는 다양한 분야에서 활용 되고 있다. 현대 전쟁에서의 공격용 UAV 운용, 재난 환경에서 인간이 식별하기 어려운 지역을 정찰 및 탐지 등의 재난 구조 UAV, 공중 기지국 활용, 그리고 도심 항공 교통(urban air mobility, UAM)과 같은 차세대 교통 수단 등으로 주목 받고 있다. 앞서 다양한 분야에서 연구되는 UAV를 운용 시, UAV 기체 자세를 제어하는 것은 굉장히 중요하다. 기존의 UAV 기체 자세 제어 방법으로는 PID 제어 방식이 대표적으로 활용되었다. 하지만, 기존 PID 제어를 활용한 UAV 기체 자세 제어 방법은 빠르게 변하는 실제 운용 환경에서 불안정하며, 특히 UAM과 같은 기체의 크기가 큰 UAV의 자세를 제어하기에는 성능이 부족하다 [1].

따라서, 본 논문에서는 강화학습 기법 중 deep deterministic policy gradient(DDPG) 알고리즘 기반 자세 제어 방법을 실제 UAM의 기체 스펙에 맞게 설계하여 시뮬레이션을 진행하고 UAM 자세 제어 안정성의 성능을 분석한다 [2].

### II. UAV 동역학 모델

본 논문에서 고려한 UAM의 구조는 회전 모터 4개가 탑재된 쿼드콥터 UAM을 고려하며 회전 동역학을 그림 1과 같이 표현한다. UAM의 위치와 자세 등 회전 동역학은 기체 중심 좌표계( $x_b, y_b, z_b$ )와 지구 중심 좌표계( $x_e, y_e, z_e$ )의 관계로 나타낸다. UAM의 회전 각도는 Euler Angle을 기준으로,  $x_b$ 를 중심으로 회전하는 각도를  $\phi$ (roll),  $y_b$ 축 중심 회전 각도를  $\theta$ (pitch),  $z_b$ 축 중심 회전 각도를  $\psi$ (psi)라 하며, 각 회전 각도별 각속도는  $\dot{\phi}, \dot{\theta}, \dot{\psi}$ 로 정의한다. 각 모터의 추력은  $T_i$ 로 정의하며 1번과 3번 모터는 시계방향, 2번과 4번 모터는 반시계 방향으로 회전함으로써 비행 추력을 생성한다. 동역학과 추력과의 관계와 Euler's 회전 방정식을 기반으로 UAM의 회전 역학 방정식( $M$ )은  $M = M_\tau + M_G + M_f$ 로 정의한다. 각 모터의 제어 토크( $M_\tau = [\tau_\phi, \tau_\theta, \tau_\psi]$ ), 각 회전 모터의 관성 모멘트

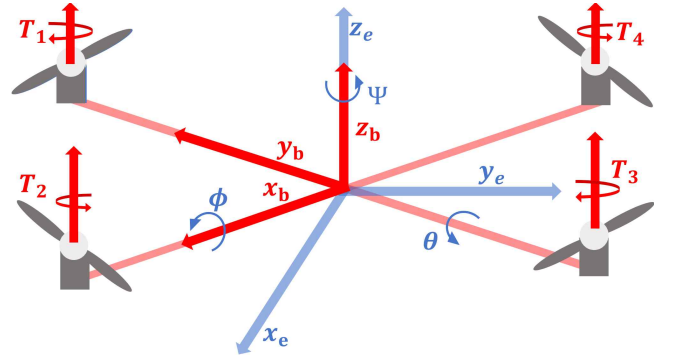


그림 1. UAM 동역학 모델

( $I_p$ )와 각 축 별 각속도, 방해 효과( $\Omega$ )로 인하여 생성된 자이로 스코프 효과( $M_G = [-I_p \dot{\theta} \Omega, I_p \dot{\phi} \Omega, 0]^T$ ), 그리고 회전 시 발생하는 항력은 회전 항력( $d_i$ )과 각속도의 관계( $M_f = [-d_\phi \dot{\phi}, -d_\theta \dot{\theta}, -d_\psi \dot{\psi}]^T$ )로 정의하며,  $M_\tau$ 는 각 모터의 추력( $T_i$ )과 모터와 UAM 중심 사이 거리 ( $L$ ), 추력과 비례한 리액션 토크( $K_\psi$ )의 관계로 수식 (1)과 같이 나타낸다.

$$M = \begin{bmatrix} \tau_\phi \\ \tau_\theta \\ \tau_\psi \end{bmatrix} = \begin{bmatrix} L(T_2 - T_4) \\ L(T_1 - T_3) \\ K_\psi(T_1 - T_3 + T_2 - T_4) \end{bmatrix}, \quad (1)$$

UAM의 비행 이동 방정식( $F_e$ ) 구조는  $F_e = RF_l + F_d + G$ 로 정의하며, 상승하는 힘( $F_l = [0, 0, T_z]$ )과 공기 역학 항력( $F_d$ ), 그리고 중력( $G$ )의 관계이다. 비행 이동 간 쿼드콥터 UAM의 회전 행렬( $R$ )은 수식 (2)와 같다.

$$R = \begin{bmatrix} C_\phi C_\theta & C_\phi S_\theta S_\psi - C_\phi C_\psi S_\theta & C_\phi C_\psi S_\theta + S_\phi S_\psi \\ S_\phi C_\theta & S_\phi S_\theta S_\psi + C_\phi C_\psi & S_\phi S_\theta S_\psi - C_\phi C_\psi \\ -S_\theta & C_\theta S_\phi & C_\theta C_\phi \end{bmatrix}, \quad (2)$$

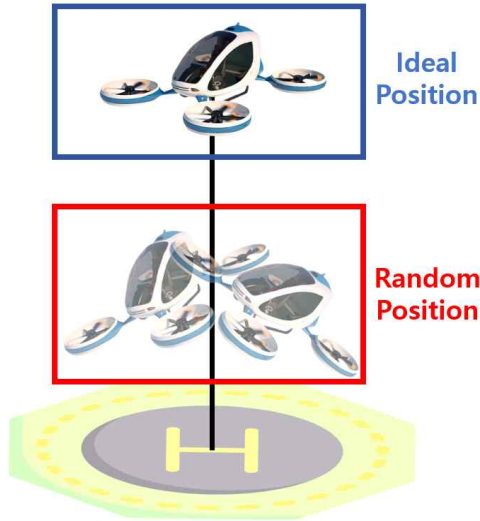


그림 2. 시스템 모델

표 1. UAM 상세 스펙

| Parameter                  | Value                    |
|----------------------------|--------------------------|
| UAM 크기 (가로, 세로, 높이)        | 2.5 m, 1.5 m, 0.535 m    |
| $L$                        | 1.0 m                    |
| $T_1, T_2, T_3, T_4$       | 400 ~ 500 N              |
| $K_\psi$                   | 0.4 N                    |
| $I_p$                      | 0.03 kg · m <sup>3</sup> |
| $d_\phi, d_\theta, d_\psi$ | 0.1, 0.1, 0.2            |
| $G$                        | 9.81 m/s <sup>2</sup>    |

### III. 시스템 모델

UAM의 비행 안정성을 실험하기 위한 시스템 모델은 그림 2와 같다. UAM의 시작 자세 좌표는 일정 범위 내에서 무작위로 설정되며 이상적인 자세  $(x_b, y_b, z_b) = (0, 0, 0)$ 를 목표로 제어를 하며, UAM의 비행 시작 지점과 비행 목표 지점은 고정되어 있다. 다음은 강화학습을 진행하기 위한 마르코프 결정 프로세스(Markov decision process, MDP)의 상태, 행동, 보상에 관한 정의다.

1) 상태(states)  $\mathcal{S} \doteq [x, y, z, x_b, y_b, z_b, \phi, \theta, \psi]$ 로 정의한다. UAM의 현재 위치를  $x, y, z$ , UAM의 현재 자세 좌표  $x_b, y_b, z_b$ , 그리고 각 축의 회전 각도  $\phi, \theta, \psi$ 로 나타낸다.

2) 행동(action)  $\mathcal{A} \doteq [T_1, T_2, T_3, T_4]$ 로 정의한다. 에이전트는 UAM의 각 모터의 추력을 조절하여 기체의 자세를 제어하며 목적지까지 자세의 안정성을 유지하며 비행하는 것을 목표로 학습한다.

3) 보상(reward)  $R$ 은 현재 UAM의 비행 위치와 목표 지점 사이의 거리, 현재 UAM의 자세와 이상적인 자세 차이인 tracking error 값을 기반으로 설계한다. 에이전트는 보상 값을 최대화하는 방향으로 학습을 진행하며 상세 보상 함수는 수식 (3)과 같이 정의한다.

$$R \doteq 1 - (\alpha * (x_p^2 + y_p^2 + z_p^2) + \beta(x_e^2 + y_e^2 + z_e^2)), \quad (3)$$

강화학습 에이전트인 UAM의 위치 좌표, 기체 자세 등의 상태 값은 실숫값의 형태로 연속적으로 변화한다. 이에 본 논문에서는 연속적으로 변화하는 환경에서 학습에 용이한 DDPG 기법을 활용하여 학습을 진행하였으며, UAM의 상세 스펙은 표 1, 강화학습 파라미터는 표2와 같다.

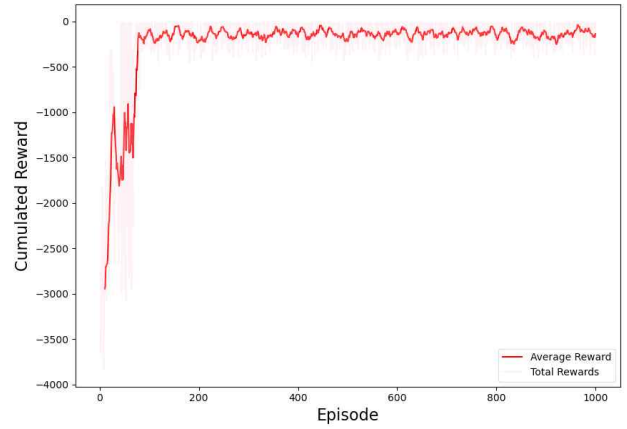


그림 3. 강화학습 결과

표 2. 강화학습 파라미터

| Parameter          | Value   |
|--------------------|---------|
| Batch size         | 128     |
| Learning rate      | 0.001   |
| Discount factor    | 0.99    |
| Replay buffer size | 100,000 |

### IV. 시뮬레이션 결과

시뮬레이션 시나리오로는 그림 2와 같이 UAM의 비행 시작 지점에서 수직으로 10m 떨어진 지점을 향해 비행하는 시나리오를 기반으로 무작위로 설정된 UAM의 불안정한 기체 자세를 안정적인 비행 자세로 회귀하기 위해 각 모터의 추력을 조절하는 방향으로 학습을 진행하였다. 시뮬레이션 결과, 약 200 에피소드 이후로 보상함수의 값이 수렴하는 것과 경향성을 기반으로 자세 제어 학습이 원활하게 이루어졌다는 것을 확인했다.

### V. 결론

본 논문에서는 모터가 4개인 쿼드콥터의 동역학을 기반으로 도심 항공 택시 UAM의 비행 안정성을 확보하기 위해 강화학습 기반 UAM 자세 제어 기법을 제안했다. 제안하는 강화학습 기법으로 다양하게 변화하는 기상 환경 혹은, 돌발 상황에서도 안정적인 비행이 가능토록 하는 강화학습 모델을 설계하였고, 추후 운용될 UAM에 제안하는 기법이 효과적인지에 대한 성능을 분석하였다.

### ACKNOWLEDGMENT

이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2021-0-00794, 3차원 공간이동통신 기술 개발)

### 참고 문헌

- [1] J. Niping, Y. Zhiweri, and K. Yang, "Operational effectiveness evaluation of the swarming UAVs combat system based on a system dynamics model," *IEEE Access*, vol. 7, pp. 25209-25224, 2019.
- [2] Y. Wang, J. Sun, H. He, and C. Sun, "Deterministic policy gradient with integral compensator for robust quadrotor control," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 50, pp. 3713-3725, 2020.