

강화학습 기반 항공기 세로 정적 안정성 제어

김규선, 박수현*, 김종현

고려대학교, 숙명여자대학교*

kingdom0545@korea.ac.kr, soohyun.park@sookmyung.ac.kr*, joongheon@korea.ac.kr

Reinforcement Learning-based Aircraft Longitudinal Static Stability Control

Gyu Seon Kim, Soohyun Park*, Joongheon Kim

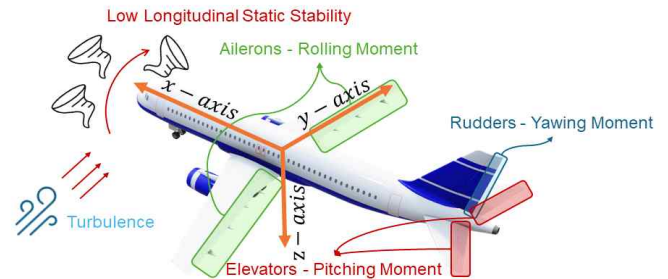
Korea Univ., Sookmyung Women's Univ.*

요약

최근 온실가스 배출량 증가로 인한 제트 기류의 풍속 변화 영향으로 난기류가 증가하고, 이로 인해 항공기(민항기)가 비행 안정성을 잃어 순간적으로 급하강하는 항공 사고가 빈번하게 발생하고 있다. 잦은 난기류의 등장으로 인한 여러 항공 사고의 발생은 강건한 항공기 세로 정적 안정성 제어의 필요성을 부각시켰다. 여러 난기류 중 항공기의 세로 방향 모멘트를 증가시키는 난기류가 특히 많이 발생하기 때문에, 3개의 축을 기반으로 한 항공기 안정성 중 특히 피칭모멘트를 기반으로 한 세로 정적 안정성이 항공기 안전에 있어 매우 중요한 요소로 인식되고 있다. 그러나 기존의 Proportional-Integral-Derivative (PID) 제어기는 고정된 파라미터를 기반으로 동작하기 때문에 변화하는 비행 환경에 적응하는 데 한계가 있다. 즉, 갑자기 발생하는 강한 난기류나 기상 변화와 같은 동적인 환경에서 PID 제어기는 최적의 성능을 유지하기 힘들다. 또한 선형 시스템을 가정하여 설계된 PID 제어기는 매우 복잡하고 비선형적인 특성을 가진 항공기의 역학시스템을 완전하게 반영할 수 없다. 이러한 문제점을 해결하기 위해 본 논문에서는 기계학습(Machine Learning, ML) 방법 중 하나인 강화학습(Reinforcement Learning, RL)을 활용하여 항공기의 세로 정적 안정성을 개선하는 새로운 접근 방식을 제안한다. 특히, 항공기의 수평꼬리날개 조종간인 엘리베이터(Elevator)의 각도를 강화학습을 통해 학습 및 제어하여 강한 난기류 상황에서도 항공기의 세로 정적 안정성을 유지한다. 훈련된 항공기 제어 모델은 가변적인 난류 조건 하에서도 항공기가 더 빠르고 정확하게 반응하여 세로 정적 안정성을 유지할 수 있도록 설계되었다. 실제 비행 환경과 유사하게 무작위한 강한 난기류가 존재하는 실험환경에서의 검증은 본 논문에서 제안된 제어 알고리즘의 우수성을 평가함과 동시에 항공기 제어 분야에 인공지능(Artificial Intelligence, AI) 기술을 적용하기 위한 새로운 길을 열어준다.

I. 서론

최근 기후 변화는 항공 산업에 예상치 못한 도전을 제시하고 있다. 특히, 온실가스 배출량의 증가가 제트 기류에 영향을 미치면서 항공기가 비행하는 동안 난기류에 의한 위험이 증가하고 있다. 이러한 난기류는 항공기의 세로 정적 안정성을 위협하며, 심각한 경우 항공기의 급격한 피치 변화를 유발하여 안전한 비행을 불가능하게 만든다. 실제로 최근 영국 런던에서 출발하여 싱가포르로 향하던 싱가포르 항공 여객기(SQ321편 Boeing 777-300ER)가 상공 고도 37,000 ft에서 심각한 난기류를 만나 제어시스템이 붕괴되면서 3분 동안 약 6,000 ft 급하강하고 방콕 수완나폼 국제공항에 비상 착륙하는 항공 사고가 발생하였다 [1]. 이때 순간적으로 중력가속도가 0.6초 동안 +0.35G에서 -1.5G로 감소했다가 다시 4초 만에 -1.5G에서 +1.5G로 바뀌었다. 짧은 시간 동안의 중력가속도 변화로 인해 승객들은 순간적으로 여객기 천장으로 솟구쳤다가 바닥으로 떨어졌고 이 과정에서 수많은 부상자와 사망자까지 발생하였다. 이 사건은 난기류에 대응하는 현재의 항공기 제어 시스템의 한계를 여실히 보여준다. 이러한 문제의 심각성에 대응하여, 항공기의 제어 시스템은 점점 더 높은 수준의 자동화와 지능화를 요구하고 있다. 기존의 Proportional-Integral-Derivative (PID) 제어기는 고정된 매개변수를 기반으로 동작하기 때문에, 빠르게 변화하는 비행 환경과 복잡한 비선형 항공 동역학 시스템에 효과적으로 대응하는 데 한계가 있다. 이는 항공기가 심각한 난기류를 만났을 때 적절히 반응하지 못하고 안정성을 잃을 가능성이 높다는 것을 의미한다. 반면, 강화학습(Reinforcement Learning, RL)은 주어진 환경에서의 상호작용을 통해 최적의 항공기 제어 방법을 스스로 학습하는 기계학습(Machine Learning, ML)의 한 형태로, 이러한 동적인 상황에 보다 유연하게 대응할 수 있는 잠재력을 지니고 있다 [2]. 이에 본 논문은 강화학습을 활용하여 항공기 세로 정적 안정성(Longitudinal Static Stability, LSS)을 개선하는



[그림 1] 항공기 각 축의 조종간과 세로 정적 안정성

강건한 항공기 자세 제어 알고리즘을 제안한다. 특히, 항공기의 수평 꼬리날개인 엘리베이터(Elevator)의 각도 조절을 학습하는 강화학습 알고리즘을 개발하여, 가변적인 난기류 상황에서도 항공기가 안정적인 비행을 유지할 수 있도록 한다. 본 논문에서는 가치 기반 방법(Value based Method)의 대표적인 알고리즘인 Deep Q-Network (DQN)과 정책 기반 방법(Policy based Method)의 대표적인 알고리즘인 Deep Deterministic Policy Gradient (DDPG)를 사용하여 강화학습 기반 항공기 세로 정적 안정성 제어 알고리즘을 설계한다. 본 논문은 실제 비행 환경과 유사하며 강한 무작위 난기류가 존재하는 현실적인 실험 환경에서 강화학습 기반 항공기 자세 제어 시스템의 성능을 검증함으로써, 항공기 제어 분야에 인공지능 기술을 적용할 수 있는 새로운 가능성을 탐구한다.

II. 1. 항공기 세로 정적 안정성 제어 시스템과 강화학습 설계

[그림 1]은 항공기의 조종간과 세로 정적 안정성을 보여준다. 항공기는 3축 운동을 하기에 각 축에 대한 조종간이 있는데 x축은 롤링모멘트로 에일러론에 의해, y축은 피칭모멘트로 엘리베이터에 의해, z축은 요잉모멘트로 러더에 의해 제어된다. 세로 정적 안정성의 경우 엘리베이터에 의해 항공기의 자세각이 제어된다. [그림 1]과 같이 난기류에 의해 항공기의 코

를 들어올리는 모멘트가 갑자기 발생하면, 항공기의 세로 정적 안정성이 떨어지는데 이를 방지하기 위해 엘리베이터 각도를 적절히 제어해야 한다. 항공기가 세로 정적 안정성을 유지하려면 다음과 같은 조건을 만족해야 하고 이는 강화학습에서 보상함수와 연결된다.

$$C_{m_{\delta_e}} = \frac{dC_m}{d\delta_e} = - \left(\frac{dC_L}{d\alpha(\Omega)} \right)_t \Xi \eta_t \frac{d\alpha_t(\Omega)}{d\delta_e} < 0 \quad (1)$$

식(1)에서 $C_{m_{\delta_e}}$, C_m , δ_e , C_L , Ξ , η_t , $\alpha_t(\Omega)$, Ω 는 각각 엘리베이터 각에 따른 피칭모멘트 계수 변화량, 피칭모멘트 계수, 엘리베이터 각, 양력계수, tail volume, tail efficiency, 받음각(Angle of Attack, AOA), 그리고 난기류의 세기를 나타내는 상수값이다. 강화학습의 관점에서 항공기의 행동은 엘리베이터 각 제어이며 보상함수는 $R = \Gamma \cdot \exp(-C_{m_{\delta_e}})$ 로 정의된다. Γ 는 엘리베이터 각에 따른 피칭모멘트 계수 변화량의 민감도를 조절하는 상수이다. 상태는 난기류의 세기, 받음각, 양력계수, 항공기의 자세각, 엘리베이터의 틸트각이며 이러한 물리량들이 신경망의 입력값이다.

II. 2. PID 제어기

PID 제어기는 자동 제어 분야에서 가장 널리 사용되는 피드백 루프 메커니즘 중 하나로 간단한 구조와 넓은 적용 범위 때문에 다양한 산업 분야에서 활용되고 있다. PID 제어기는 시스템의 출력을 원하는 설정값(set point)에 맞추기 위해 입력을 조절한다. PID 제어기는 이름에서 알 수 있듯, 세 가지 주요 구성 요소: 비례(Proportional), 적분(Integral), 미분(Derivative) 요소를 통합하여 시스템의 동적 성능을 개선하고 오차를 최소화하는 것을 목표로 하며, 이는 수학적으로 다음과 같이 표현된다.

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{d}{dt} e(t) \quad (2)$$

식 (2)에서 $u(t)$, K_p , K_i , K_d , $e(t)$ 는 각각 시간 t 에서의 출력, 비례 계인, 적분 계인, 미분 계인, 오차를 의미한다.

II. 3. 강화학습과 DQN 알고리즘

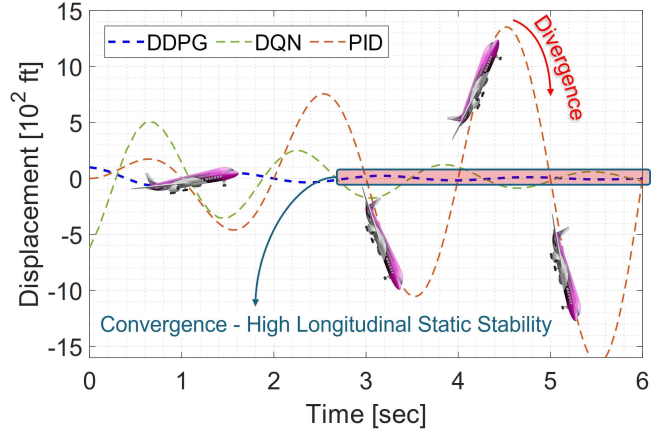
강화학습은 에이전트가 환경과의 상호 작용을 통해 학습하여 누적 보상을 최대화하는 기계학습 내의 방법론이다. 환경을 탐색하는 에이전트는 다양한 상태에서 행동의 가치를 학습하고 이러한 값을 최적화하기 위해 지속적으로 행동을 조정한다. 에이전트와 환경 간의 이러한 반복적인 상호 작용이 강화학습의 핵심이며, 이는 항공기 자세 각 제어와 같이 매우 동적인 환경에서의 순차적 의사 결정 문제에 특히 적합하다. 강화학습 알고리즘 중 DQN은 Bellman 최적 방정식을 학습 타겟으로 간주하며, 학습에 사용되는 손실 함수는 식 (3)로 정의된다.

$$L(\theta)' = E[(R_t + \gamma \max_{a_{t+1}} Q_{\theta^-}(s_{t+1}, a_{t+1}; \theta^-) - Q_{\theta}(s_t, a_t; \theta))^2] \quad (3)$$

식 (3)에서 $L(\theta)'$, R_t , γ , $Q_{\theta^-}(s_{t+1}, a_{t+1}; \theta^-)$ 그리고 $Q_{\theta}(s_t, a_t; \theta)$ 는 각각 손실함수, 보상 함수, 감쇠 인자, 타겟 네트워크에 의해 추정된 Q값, 그리고 에이전트가 얻은 Q값이다. 여기서 θ 는 신경망 파라미터이다.

II. 4. DDPG 알고리즘

DDPG는 연속적인 행동 공간을 가진 환경에서 효과적으로 작동하는 강화학습 알고리즘이다. DDPG는 Actor-Critic 방법론을 기반으로 하며, 결정론적 정책을 사용하여 최적의 행동을 직접 학습한다. 이는 높은 차원의 행동 공간에서도 효과적으로 학습하며, 특히 복잡한 시뮬레이션 환경에서의 항공기 제어와 같은 문제에 적합하다. DDPG는 두 개의 주요 네트워크, 즉 Actor와 Critic 네트워크를 사용하는데, Actor 네트워크는 주어진 상태에 대해 최적의 행동을 결정하는 정책을 출력하고 Critic 네트워크는 Actor가 선택한 행동의 가치를 평가한다. 여기서 Actor 네트워크의 그래디언트와 Critic 네트워크의 그래디언트는 각각 식 (4), (5)로 표현된다.



[그림 2] 각 알고리즘별 항공기 세로 정적 안정성 추이

$$\nabla_{\theta^{\mu}} J = E[\nabla_a Q(s, a | \theta^Q)|_{a=\mu(s)} \cdot \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu})] \quad (4)$$

$$\nabla_{\theta^{\mu}} L = E[\nabla_{\theta^Q} Q(s, a | \theta^Q) \cdot (Q(s, a | \theta^Q) - y)] \quad (5)$$

식(4)와 식(5)에서 J , $Q(s, a | \theta^Q)$, $\mu(s | \theta^{\mu})$, L , 그리고 y 는 각각 Actor의 목적함수, Critic 네트워크가 계산하는 Q-값, Actor 네트워크가 주어진 상태에 대해 선택한 행동, Critic의 손실함수, 그리고 타겟 Q값을 의미한다. DDPG는 이를 기반으로 최적의 정책을 학습해 나간다.

III. 성능평가

[그림 2]는 제안된 알고리즘과 벤치마크별 항공기 세로 정적 안정성 추이를 보여준다. x축은 시간[s], y축은 세로 방향 변위[10² ft]이다. PID 제어기는 비행 중 난기류를 만나면 정적 안정성을 회복하지 못하고 발산하는 모습을 보여준다. 그에 반해 강화학습 기반 세로 정적 안정성 제어기는 난기류를 만나도 안정적으로 다시 자세를 유지한다. 강화학습 알고리즘 중에서도 DDPG 기반의 제어기가 DQN 기반의 제어기보다 더 안정적으로 항공기의 자세각을 제어하는 것을 알 수 있다. 이는 불연속적인 행동(엘리베이터 각도값)을 학습하는 DQN과 달리 연속적인 엘리베이터 각도값을 학습할 수 있는 DDPG의 장점으로 인해 나온 결과로 해석될 수 있다.

IV. 결론

본 논문에서는 강화학습을 통해 항공기 세로 정적 안정성 제어 문제를 해결하였으며, 강화학습 기반 항공기 제어기가 기존의 PID 제어기보다 난기류가 심한 비행 환경에 효과적으로 대응할 수 있음이 입증되었다. 특히, DDPG 기반의 제어기가 DQN 기반의 제어기보다 난기류가 심한 조건에서 항공기의 피칭모멘트를 더 안정적으로 제어함으로써 비행 안전성을 크게 향상시킬 수 있는 가능성을 보여주었다.

ACKNOWLEDGMENT

This work was supported by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00467, Intelligent 6G Wireless Access System)

참고 문헌

- [1] Jacopo Prisco, "Climate change could be about to make flight turbulence a lot worse", CNN, Wed May 22, 2024.
- [2] G. S. Kim, J. Chung and S. Park, "Realizing Stabilized Landing for Computation-Limited Reusable Rockets: A Quantum Reinforcement Learning Approach," *IEEE Transactions on Vehicular Technology*, pp. 1-6, 2024 (Early Access).