

# 다기종 무인 플랫폼의 확장성 확보를 위한 행동 추상화 기반의 피지컬 AI 제어 시스템 연구

김정현 , 신수용

국립금오공과대학교

kjh454311@kumoh.ac.kr, wdragon@kumoh.ac.kr

## A Study on Physical AI Control System based on Action Abstraction for Scalability of Heterogeneous Unmanned Platforms

Jung Hyun Kim , Soo Young Shin  
Kumoh National Institute of Technology

### 요약

본 연구는 다수·다기종 무인 플랫폼(HMRS)의 효율적인 통합 운용을 위해, 피지컬 AI 기반의 계층적 제어 아키텍처를 제안한다. 기존의 제어 방식은 플랫폼의 이기종성으로 인해 제어 복잡도가 높고 확장이 제한적이라는 한계를 가진다. 이를 해결하기 위해 본 연구는 '전술 판단'과 '물리적 실행'을 구조적으로 분리하고, 이를 '행동 추상화(Action Abstraction)' 계층으로 연결하여 상위 전술 명령을 기종별 구체적 행동으로 자동 변환하는 시스템을 설계하였다. 제안된 아키텍처는 하드웨어 변경 시에도 소프트웨어의 전면 수정 없이 유연한 확장이 가능하며, MLOps 파이프라인의 통합을 통해 비정형 환경에서도 지속 가능한 운용성을 보장한다.

### 1. 서론

최근 로봇틱스 연구는 단일 고성능 로봇 개발을 넘어, 공중(UAV)·지상(UGV)·해상(USV) 등 다양한 무인 플랫폼을 하나의 유기적인 체계로 통합하는 이기종 다중 로봇 시스템(Heterogeneous Multi-Robot Systems, HMRS)으로 진화하고 있다[1]. 특히 현대전 및 대규모 재난 현장에서는 저비용의 다수 무인 플랫폼을 활용하여, 일부가 소실되더라도 전체 시스템의 임무 수행 능력을 유지하는 유연한 군집 운용이 핵심 경쟁력으로 부상했다[2].

그러나 플랫폼의 종류와 수가 증가함에 따라, 기존의 '1인 1기' 원격 제어나 단순 태스크 할당(Task Allocation) 방식은 한계에 직면했다[3]. 이기종 플랫폼 간의 상이한 통신 프로토콜과 제어 방식은 시스템의 통합 복잡도를 기하급수적으로 증가시키며, 새로운 기종 추가 시 전체 시스템을 재설계해야 하는 확장성(Scalability) 문제를 야기한다[4].

이에 본 연구는 거대언어모델(LLM)과 같은 피지컬 AI 기술을 응용하여 [5], 시스템의 '두뇌(판단)'와 '몸통(실행)'을 구조적으로 분리하는 계층적 아키텍처를 제안한다. 기존의 Google SayCan과 같은 연구가 단일 로봇의 행동 생성에 집중했다면[6], 본 연구는 이를 다기종 시스템으로 확장하여 \*\*'행동 추상화(Action Abstraction)\*\*\*를 통해 이기종 간의 물리적 장벽을 소프트웨어적으로 극복하는 데 목적이 있다.

### 2. 관련 연구

#### 2.1 이기종 무인 체계의 군집 운용 및 소프트웨어 정의 전장

전통적인 무인 체계 연구가 고성능 단일 플랫폼의 자율성 향상에 집중했다면, 최근의 국방 및 로봇틱스 트렌드는 저비용·다수·다기종 플랫폼(HMRS)의 유기적인 통합 운용으로 패러다임이 전환되고 있다[1]. 특히 미 국방부의 '레플리케이터(Replicator)' 프로젝트는 수천 대의 자율 무인 체계를 신속하게 생산 및 배치하여, 일부 자산이 소실되더라도 전체 시스템의 지능과 작전 능력이 유지되는 유연한 전력 구성을 목표로 한다.

이러한 물량 중심의 하드웨어 운용을 가능하게 하는 핵심은 소프트웨어에 있다. 기존의 중앙 집중식 제어는 단일 실패 지점(Single Point of Failure)

문제와 통신 대역폭의 한계로 대규모 확장에 취약하다[2, 3]. 이에 대응하여 안두릴(Anduril)의 '라티스(Lattice)'와 같은 시스템은 개별 로봇의 세부 제어보다는, 임무 단위의 지휘를 통해 이기종 플랫폼들이 자율적으로 역할을 분담하고 협업하는 통합 운영 체계를 제시하고 있다. 중국의 '늑대 로봇' 사례 또한 경찰-타격-보급 로봇이 군집을 이루어 유기적으로 협동함으로써, 실전 환경에서 인명 피해를 최소화하고 효율성을 극대화하는 군집 로봇 기술의 실증 사례를 보여준다.

#### 2.2 피지컬 AI와 거대언어모델(LLM)의 로봇틱스 적용

복잡한 비정형 환경에서 다수 로봇을 제어하기 위해서는 단순한 규칙 기반(Rule-based) 로직을 넘어선 고도의 상황 인지 능력이 요구된다. 최근 거대언어모델(LLM)과 멀티모달 AI를 포함한 피지컬 AI(Physical AI) 기술은 로봇이 자연어 명령을 이해하고 상황의 맥락(Context)을 파악하여 다음 행동을 능동적으로 추론하는 능력을 제공한다.

Google의 'SayCan'과 같은 연구[6]는 LLM의 언어적 추론 능력을 로봇의 행동 가능성(Affordance)과 결합하려는 시도였으나, 이는 주로 단일 로봇의 행동 생성에 국한되었다. 다기종 환경에서는 각 플랫폼이 수행할 수 있는 행동의 종류와 물리적 제약이 서로 다르므로, 상위의 지능(Brain)과 하위의 실행(Body)을 연결하는 새로운 아키텍처가 필요하다.

#### 2.3 행동 추상화 및 아키텍처의 필요성

기존 연구들은 주로 이기종 플랫폼 간의 데이터 통신 프로토콜 통합이나 미들웨어 수준의 연결에 집중하였다[4]. 그러나 플랫폼의 종류가 늘어날수록 시스템 복잡도가 급증하는 문제를 해결하기 위해서는 '판단'과 '실행'을 명확히 분리하는 접근이 필수적이다.

본 연구는 선행 연구의 한계를 극복하기 위해 '행동 추상화(Action Abstraction)' 개념을 도입한다. 이는 전술적 판단을 플랫폼과 무관한 '의미 단위의 행동 요청'으로 정의하고, 이를 중계 계층에서 각 플랫폼에 최적화된 실행 명령으로 변환한다. 이러한 구조는 새로운 플랫폼이 추가되더라도 상위 판단 구조를 수정할 필요가 없어 시스템의 확장성(Scalability)을 보장하며, MLOps 루프를 통해 운용 데이터를 학습하여 지속적으로 진화할 수 있다는 점에서 기존 정적 제어 시스템과 차별화된다.

### 3. 제안하는 시스템 아키텍처

본 연구는 다수·다기종 무인 플랫폼의 복잡성을 해결하고 확장성을 확보하기 위해, 인간의 인지-행동 과정을 모사한 3단계 계층 구조(Three-Layered Architecture)를 제안한다. 이 구조는 판단(Decision)과 실행(Execution)을 명확히 분리하여, 상위 시스템은 "무엇을 할 것인가(What)"에 집중하고 하위 시스템은 "어떻게 할 것인가(How)"를 전담하는 것을 원칙으로 한다.

#### 3.1 전술 판단 계층 (Tactical Decision Layer)

최상단에 위치한 전술 판단 계층은 피지컬 AI(Physical AI) 기반의 시스템 두뇌 역할을 수행한다. 이 계층은 개별 센서 데이터나 로봇의 좌표 정보보다는 전장 상황의 '맥락(Context)'을 이해하는 데 집중한다.

맥락 기반 의사결정: "현재 임무가 어느 단계인가", "가용한 자산은 무엇인가"와 같은 고수준의 정보를 융합하여 상황을 인지한다.

전술적 행동 도출: 구체적인 로봇 제어 명령(예: 모터 속도, 타각)이 아닌, "감시 공백 보완", "의심 객체 추적"과 같은 전술적 의미를 가진 행동 요구 사항(Action Request)을 산출한다. 이 판단은 하드웨어 플랫폼의 종류와 관계없이 동일하게 내려질 수 있는 공통의 판단 구조이다.

#### 3.2 임무 오케스트레이션 계층 (Mission Orchestration Layer)

본 아키텍처의 핵심 기여점인 중간 계층으로, 전술 판단 계층과 로컬 실행 계층 사이에서 '행동 추상화(Action Abstraction)' 및 중재자(Mediator) 역할을 수행한다.

행동 번역(Translation): 상위 계층의 추상적인 행동 요청을 각 플랫폼이 이해할 수 있는 구체적인 실행 명령으로 변환한다. 예를 들어, 동일한 "추적 유지" 명령이라도 드론에게는 '상공 비행 및 카메라 줌인'으로, 지상 로봇에게는 '경로 확보 및 접근'으로 서로 다르게 매핑된다.

확장성(Scalability) 보장: 새로운 기종의 로봇이 추가되더라도, 이 계층에서 해당 로봇이 수행 가능한 행동(Capabilities)만 정의하면 된다. 상위의 전술 판단 로직을 수정할 필요 없이 기존 자산 하나로 자연스럽게 통합할 수 있어 시스템의 유연성을 극대화한다.

#### 3.3 로컬 실행 계층 (Local Execution Layer)

최하단에 위치한 로컬 실행 계층은 각 무인 플랫폼(UAV, UGV, USV 등)의 엣지(Edge) 단에서 물리적 제어를 담당한다.

플랫폼 최적화: 센서 데이터 처리, 위치 추정(Localization), 자세 제어 등은 각 플랫폼의 동역학적 특성에 맞춰 최적화되어야 하며, 중앙 시스템은 이러한 내부 구현(Internal Implementation)을 알 필요가 없다.

안전 및 생존성: 충돌 회피나 통신 두절 시의 자동 회귀(Return to Home)와 같은 즉각적인 생존 본능은 상위 계층의 판단 대기 없이 로컬에서 독립적으로 수행하여 안전성을 보장한다.

### 4. 시스템 운용 메커니즘 및 MLOps 전략 (System Operation Mechanism & MLOps Strategy)

#### 4.1 엣지-클라우드 협업 운용

전술 환경의 통신 불안정성을 극복하기 위해 본 시스템은 이원화된 추론 구조를 갖는다. 즉각적인 반응(장애물 회피 등)은 엣지 디바이스에서 처리하고, 고수준의 전술 판단은 중앙 서버에서 수행한다. 통신이 단절될 경우 로컬 계층은 독립 모드로 전환되어 플랫폼의 손실을 방지한다.

#### 4.2 MLOps 기반의 지속적 진화

단순한 제어를 넘어 시스템의 지속 가능성을 확보하기 위해, 본 연구는 MLOps 루프를 아키텍처에 통합한다.

데이터 수집: 현장에서의 센서 데이터 및 AI 판단에 대한 피드백(성공/실패) 수집.

재학습: 취약 구간 데이터(Corner Case)를 중심으로 모델 자동 재학습.

점진적 배포: 검증된 모델을 엣지 및 중앙 서버에 업데이트. 이러한 순환 구조는 시간이 지날수록 시스템이 전장 환경에 적응하며 능력이 고도화되는 결과를 낳는다.

#### 5.1 결론

본 연구는 다수·다기종 무인 플랫폼의 효율적인 통합 운용을 위해 행동 추상화 기반의 피지컬 AI 제어 시스템을 제안하였다. 제안된 아키텍처는 판단과 실행을 분리함으로써 하드웨어 의존성을 제거하고 시스템의 확장성을 획기적으로 향상시켰다. 향후 연구로는 이기종 로봇 간의 표준 행동 프로토콜 정의와 실제 필드 테스트를 통한 통신 지연 및 제어 안정성 검증이 수행될 예정이다.

### ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2025-00553810, 50%) This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2025-RS-2024-00437190, 50%) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

[1] Y. Rizk, M. Awad, and E. W. Tunstel, "Cooperative heterogeneous multi-robot systems: A survey," ACM Computing Surveys (CSUR), vol. 52, no. 2, pp. 1-31, 2019.

[2] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: a review from the swarm engineering perspective," Swarm Intelligence, vol. 7, pp. 1-41, 2013.

[3] A. Khamis, A. Hussein, and A. Elmogy, "Multi-robot task allocation: A review of the state-of-the-art," Cooperative Robots and Sensor Networks, pp. 31-51, 2015.

[4] D. Tardioli et al., "Ground robotics in complex environments: Patrolling and monitoring," Journal of Field Robotics, vol. 36, no. 1, 2019.

[5] S. Vemprala et al., "ChatGPT for Robotics: Design Principles and Model Abilities," Microsoft Technical Report, 2023.

[6] M. Ahn et al., "Do As I Can, Not As I Say: Grounding Language in Robotic Affordances," arXiv preprint arXiv:2204.01691, 2022.