

GCS 발표

01/

논문 1

UAVS MEET AGENTIC AI: A MULTIDOMAIN SURVEY OF AUTONOMOUS AERIAL INTELLIGENCE AND AGENTIC UAVS

[Q

© Ranjan Sapkota¹ Konstantinos I. Roumeliotis² © Manoj Karkee¹
¹Cornell University, Department of Biological and Environmental Engineering, Ithaca, NY 14853, USA rs2672@cornell.edu, mk2684@cornell.edu
²Department of Informatics and Telecommunications, University of the Peloponnese, Tripoli, Greece

June 13, 2025 [Q

ABSTRACT [S

UAVs, 에이전트 AI를 만나다: 자율 항공 지능 및 에이전트 UAV의 다중 도메인 조사

서베이성 논문

01/

ABSTRACT [Q

Agentic UAVs represent a new frontier in autonomous aerial intelligence, integrating perception, decision-making, memory, and collaborative planning to operate adaptively in complex, real-world environments. Driven by recent advances in Agentic AI, these systems surpass traditional UAVs by exhibiting goal-driven behavior, contextual reasoning, and interactive autonomy. We provide a comprehensive foundation for understanding the architectural components and enabling technologies that distinguish Agentic UAVs from traditional autonomous UAVs. Furthermore, a detailed comparative analysis highlights advancements in autonomy, learning, and mission flexibility. This study explores seven high-impact application domains precision agriculture, construction & mining, disaster response, environmental monitoring, infrastructure inspection, logistics, security, and wildlife conservation, illustrating the broad societal value of agentic aerial intelligence. Furthermore, we identify key challenges in technical constraints, regulatory limitations, and data-model reliability, and we present emerging solutions across hardware innovation, learning architectures, and human-AI interaction. Finally, a future roadmap is proposed, outlining pathways toward self-evolving aerial ecosystems, system-level collaboration, and sustainable, equitable deployments. This survey establishes a foundational framework for the future development, deployment, and governance of agentic aerial systems across diverse societal and industrial domains.

Agentic UAV 정의:

인지(perception), 의사결정(decision-making), 기억(memory), 협력적 계획 (collaborative planning)을 통합한 자율 항공 지능의 새로운 영역.

특징 및 능력:

목표 지향적 행동(goal-driven behavior) 맥락적 추론(contextual reasoning) 상호작용적 자율성(interactive autonomy)

핵심 (Core Message)

Agentic UAV는 단순 자동화된 UAV를 넘어 인지-의사결정-기억-계획-협업 능력을 통합한 지능형 에이전트로 진화 중 정밀 농업(precision agriculture)
건설 및 채굴(construction & mining)
재난 대응(disaster response)
분야에서 활용 가능

에이전틱 UAV란

2 Foundations of Agentic UAVs |

2.1 Architectural Components [Q

The architecture of an agentic UAV (Figure 2a) is fundamentally organized; core layers: perception, cognition, control, and communication. Together, t act, and interact autonomously in dynamic environments. This layered frame (VLA) paradigm, wherein multimodal inputs are semantically interpreted a shows the architectural comparison between traditional and Agentic UAV:

1. Perception Layer: Multimodal Environmental Understanding

The perception layer is responsible for acquiring and preprocessing senso sensor suite typically includes RGB and multispectral cameras, LiDAR, t (IMUs), and barometers [49, 50]. Let $s_t \in \mathbb{R}^n$ denote the sensor input veraw data s_t to a semantic representation o_t via a function:

$$\mathbf{o}_t = \Phi(\mathbf{s}_t)$$

where $\Phi(\cdot)$ may be realized by convolutional neural networks (CNNs), transegmentation, detection, or affordance estimation. For instance, detecting c survivors in a disaster zone are both outputs of Φ .

2. Cognition Layer: Goal-Oriented Reasoning [Q

The cognition layer transforms observations into decisions [51, 52, 53, 54] planning [55, 56, 11]. Given a goal g and observation o_t , the cognitive po

$$a_t = \pi(g, \mathbf{o}_t)$$

4가지 핵심 계층

Perception (인지): 다양한 센서(RGB, 열화상, LiDAR 등) 데이터를 수집하고 AI 모델로 의미론적 이해 (예: 작물 스트레스 구역 감지)

Cognition (사고/판단) : 추론하고, 작업을 분해하며, 계획을 수립, 강화 학습 (RL) 등의 기법을 통해 최적의 행동 정책을 결정합니다. (예: 작물 병해 대처 전략 선택)

Control (제어): 실행 가능한 비행 궤적 및 구동 명령으로 변환하고, 실시간 피드 백으로 적응합니다. (예: 장애물 회피 비행)

Communication (통신) : 다른 UAV, 지상 시스템, 클라우드와 정보를 교환하고 임무를 조율합니다. (예: swarm(군집) 비행 시 정보 공유)



에이전틱 UAV란

2.2 Key Enabling Technologies

The emergence of agentic UAVs is underpinned by a convergence of key enabling technologies as that empower these aerial systems to operate autonomously, interpret complex environments, an decisions in real time. Four core technologies form the foundation of this transformation: edge Al multimodal sensing, and vision-language models.

1. Edge AI and Onboard Computing

Traditional UAVs offload data for post-processing, limiting their ability to respond in real time [6] In contrast, agentic UAVs integrate edge AI modules compact neural inference engines deployed of such as NVIDIA Jetson, Intel Movidius, or Apple Neural Engine [23, 21]. These processors s inference on the fly, enabling UAVs to perform tasks such as semantic segmentation, object reconfiguration in situ. The inference function $F_{\text{edge}}(\mathbf{s}_t)$ maps sensor input \mathbf{s}_t to decisions a_t with connection:

$$a_t = F_{\text{edge}}(\mathbf{s}_t)$$
 subject to $\tau_c < \delta$

where τ_c is the computation latency and δ is the real-time threshold (e.g., 100 ms). Efficient edge c for dynamic operations such as obstacle avoidance, anomaly tracking, and adaptive spraying.

을 수행할 수 있도록 지원하며, 클라우드 연결에 의존하지 않고 센서 입력 s_{t} 를 결정 a_{t} 로 매핑하는 추론 함수 $F_{edge}(s_t)$ 를 사용합니다.

RGB 이미지는 가시적인 구조를 포착하고, 멀티스펙트럼 센서는 식물 스트레스를 감지하 며, LiDAR는 3D 재구성을 가능하게 하고, 열화상 카메라는 열 이상 징후를 탐지합니다. 각 센서 모달리티는 다른 채널 C_i 를 제공하며, 이들의 융합 F는 통합된 관측 텐서 $o_t = F(C_1, C_2, \ldots, C_n)$ 를 형성합니다.

이 모델들은 이미지 입력 I와 언어 토큰 T를 공동으로 공유된 잠재 공간 z=f(I,T)로 인코딩하며, 여기서 시맨틱 동작이 파생됩니다.

예를 들어, VLM이 활성화된 UAV는 "손상된 태양광 패널을 검사해라" 또는 "강 근처의 저 지대 식생 지역을 찾아라"와 같은 명령에 응답할 수 있습니다.

2.2 Key Enabling Technologies (핵심 가능 기술)

Edge AI and Onboard Computing (엣지 AI 및 온보드 컴퓨팅): 데이터를 클 라우드로 보내지 않고 드론 자체에서 실시간으로 AI 추론을 수행하는 기술

Multimodal Sensing (멀티모달 센싱): 여러 종류의 센서(RGB, 멀티스펙트럴, LiDAR, 열화상 등)에서 얻은 데이터를 융합하여 환경에 대한 포괄적이고 견고한 이해를 제공

Vision-Language Models (VLMs) (시각-언어 모델): UAV가 이미지와 자연 어 명령을 동시에 이해하고 실행할 수 있도록 합니다.

2.3 Comparison with Traditional UAVs (기존 UAV와의 비교)

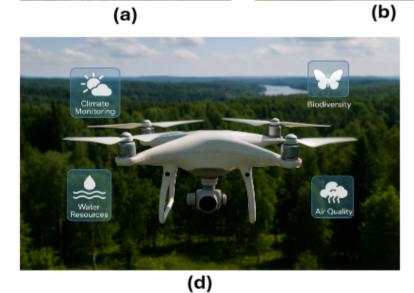
미리 프로그래밍된 임무 계획, GPS 웨이포인트 기반 비행, 사후 데이터 분석, 제 한된 자율성(Level 1-2), 고정된 작업 수행 능력등의 한계점을 동적 목표 계획, 정책 학습, 실시간 재계획 등을 통해서 극복함

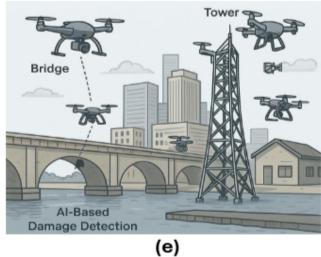
멀티 도메인

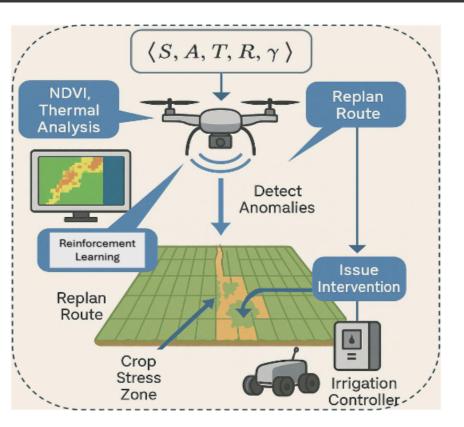












강화 학습을 통한 최적 행동 정책 (Optimal Action Policy via Reinforcement Learning)

- a) 재난 대응 및 생존자 탐지 (Disaster Response & Survivor Detection)
 - b) 홍수 및 산불 모니터링 (Flood and Wildfire Monitoring)
 - c) 서로 정보를 공유하는 스웜 코디네이션 (Swarm Coordination)
 - d) 환경 모니터링 (Environmental Monitoring)
 - e) 도시 인프라 점검 (Urban Infrastructure Inspection)

한계 및 과제



기술 및 운영 제약 (Technical and Operational Constraints)

대부분의 소형 및 중형 UAV는 리튬 폴리머(LiPo) 또는 리튬 이온 배터리를 사용하며, 이는 표준 페이로드 조건에서 20~45분 정도의 제한된 비행 시간을 제공합니다

RGB, 열화상, LiDAR, 레이더, 음향 등 여러 센서를 필요로 합니다. 그러나 추가 센서는 무게, 전력 소비, 열 발생을 증가시켜 비행 시간과 기동성에 직접적인 영향을 미칩니다

GPS 제한 또는 혼잡한 환경에서의 실시간 내비게이션 및 위치 파악에 노이즈

다중 센서 융합 및 동기화: Agentic UAV는 다양한 모달리티(시각, 열화상, 깊이, 관성 측정 장비등)의 정보를 결합하여 환경에 대한 의미론적 이해를 추론 그러나 센서들은 서로 다른 시간 및 공간 해상도로 작동하며 비선형적인 노이즈 특성

온보드 컴퓨팅 제한: 소형 UAV, 특히 엣지에서 작동하는 UAV의 경우 온보드 컴 퓨팅 성능에 상당한 제약

한계 및 과제



기술 및 운영 제약 (Technical and Operational Constraints)

대부분의 소형 및 중형 UAV는 리튬 폴리머(LiPo) 또는 리튬 이온 배터리를 사용하며, 이는 표준 페이로드 조건에서 20~45분 정도의 제한된 비행 시간을 제공합니다

RGB, 열화상, LiDAR, 레이더, 음향 등 여러 센서를 필요로 합니다. 그러나 추가 센서는 무게, 전력 소비, 열 발생을 증가시켜 비행 시간과 기동성에 직접적인 영향을 미칩니다

GPS 제한 또는 혼잡한 환경에서의 실시간 내비게이션 및 위치 파악에 노이즈

다중 센서 융합 및 동기화: Agentic UAV는 다양한 모달리티(시각, 열화상, 깊이, 관성 측정 장비등)의 정보를 결합하여 환경에 대한 의미론적 이해를 추론 그러나 센서들은 서로 다른 시간 및 공간 해상도로 작동하며 비선형적인 노이즈 특성

온보드 컴퓨팅 제한: 소형 UAV, 특히 엣지에서 작동하는 UAV의 경우 온보드 컴 퓨팅 성능에 상당한 제약

미래 로드맵

- 완전한 에이전트형 항공 생태계(Fully Agentic Aerial Ecosystems 지향)
 - 자기 진화하는 지능(Self-evolving Intelligence): 현재의 UAV는 정의된 제약 내에서 학습하고 적응하지만, 미래의 에이전트형 UAV는 새로운 기술을 자율적으로 습득하고, 변화하는 환경에 적응하며, 반복적인 평가를 통해 임무 정책을 개선할 수 있는 평생 학습(lifelong learning) 아키텍처를 통합할 것입니다. 이는 메타 학습(meta-learning), 세계 모델(world modeling), 자율 커리큘럼 생성(autonomous curriculum generation) 등의 발전을 활용합니다.
 - 성찰적 제어 아키텍처(Reflective Control Architectures): UAV가 외부 세계뿐만 아니라 자신의 목표, 행동, 한계에 대해서도 추론할 수 있도록 하는 시스템을 개발합니다. 이를 통해 현재 전략이 부적절하다고 판단될 경우, 실패 원인을 진단하고 실시간으로 계획 모듈을 재구성할 수 있습니다. 이는 특히 재난 대응이나 방위 등 불확실성이 높은 고위험 애플리케이션에서 중요합니다.
 - 이질적 에이전트 전반의 자율성 연속체 통합(Integration across the Autonomy Continuum): 개별 UAV에만 의존하는 것이 아니라, 공중, 지상, 수중 영역에서 작동하는 이질적인 플릿(fleet)이 유동적인 작업 할당과 공유된 의미론적 표현을 통해 합력하는 생태계를 구축합니다. 이 에이전트들은 임무의 중요도, 운용자 가용성, 실시간 불확실성에 따라 감독 제어(supervised control)에서 완전 자율(fully autonomous)로 자율성 연속체(autonomy continuum)를 따라 동적으로 전환될 것입니다.

미래 로드맵

- 완전한 에이전트형 항공 생태계(Fully Agentic Aerial Eco:
 - 자기 진화하는 지능(Self-evolving Intelligence): 현
 에이전트형 UAV는 새로운 기술을 자율적으로 습득하고
 개선할 수 있는 평생 학습(lifelong learning) 아키텍처를
 델(world modeling), 자율 커리큘럼 생성(autonomogeneous)
 - 성찰적 제어 아키텍처(Reflective Control Architect)
 에 대해서도 추론할 수 있도록 하는 시스템을 개발합니다
 을 진단하고 실시간으로 계획 모듈을 재구성할 수 있습니
 애플리케이션에서 중요합니다.
 - 이질적 에이전트 전반의 자율성 연속체 통합(Integratic 존하는 것이 아니라, 공중, 지상, 수중 영역에서 작동하는 적 표현을 통해 협력하는 생태계를 구축합니다. 이 에이? 라 감독 제어(supervised control)에서 완전 자율(full) 따라 동적으로 전환될 것입니다.

시스템 통합 및 협업 지능(System Integration and Collaborative Intelligence)

- UAV-UAV 협업 프레임워크(UAV-UAV collaboration frameworks): 미래의 에이전트형 UAV는 중앙 제어 없이 작업 위임, 행동 동기화, 그리고 예측 불가능한 협업(emergent coordination)이 가능한 동적 다중 에이전트 팀 (dynamic multi-agent teams)으로 작동할 것입니다. 스웜(swarm) 아키텍처는 분산 통신과 온보드 추론(onboard reasoning)을 활용하여 협업 매평, 적응형 커버리지(adaptive coverage), 분산 센상(distributed sensing)과 같은 작업을 수행합니다.
- V2X 통신(Vehicle-to-Everything Communication): UAV-UAV (V2V), UAV-지상국(V2G), UAV-인프라(V2I)
 프로토콜을 포함하는 V2X 통신은 실시간 데이터 교환, 교통 조정, 위험 완화를 가능하게 할 것입니다.
- IoT 및 엣지-클라우드 오케스트레이션(IoT and Edge-Cloud Orchestration): 에이전트형 UAV는 환경 센서, 고정 로봇, 스마트 자산과 실시간으로 통합되는 모바일 엣지 노드(mobile edge nodes) 역할을 하여 센서 네트워크 전반에 서 데이터를 수집, 융합, 전송할 것입니다.
- 디지털 트윈 기술(Digital Twin Technologies): 물리적 UAV 환경과 계산적 복제본 간의 지속적인 동기화를 가능하게 하여 예측 유지 보수, 운영 최적화, 시나리오 테스트를 지원할 것입니다.
- 항공-지상 통합(Air-Ground Integration): 자율 차량 및 지상 로봇과의 통합을 통해 물류, 농업 및 산업 자동화에서 엔드-투-엔드(end-to-end) 자율성을 지원할 것입니다.