UAV Flight Log Generation and High-Level Mission Planning Using Vision-Language Models

Author: Tsung-Yen Yu, Advisor: Dr. Wen-Hung Liao
National Chengchi University, Taipei, Taiwan

Introduction

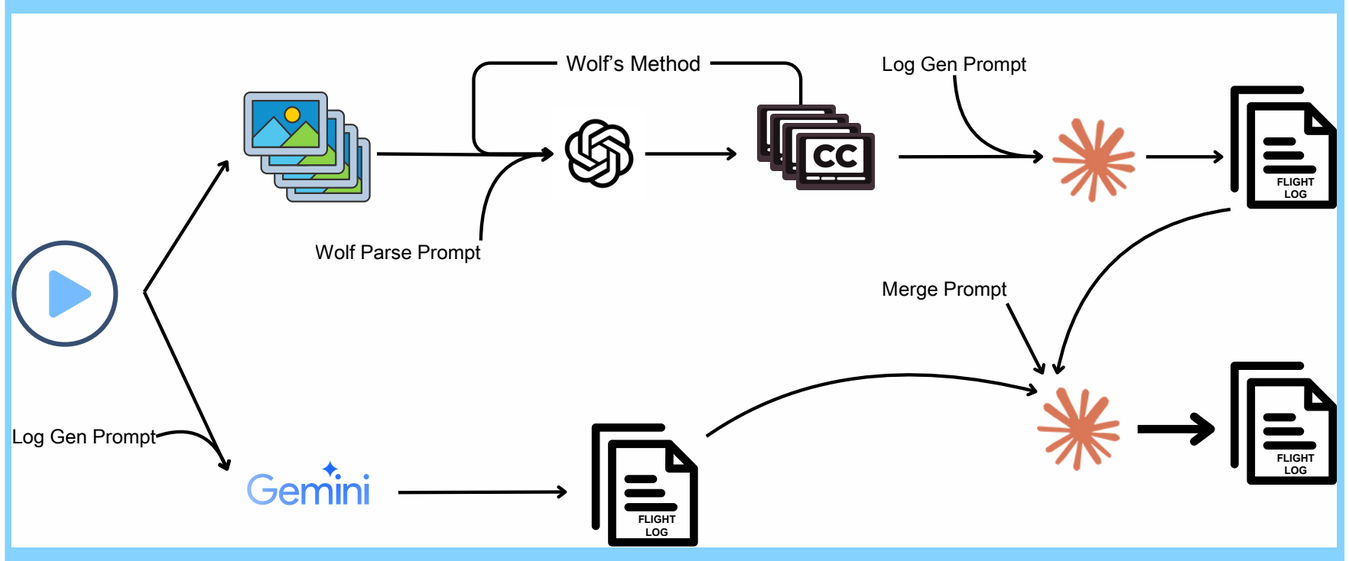
- The rapid rise of AI in computer vision is accelerating the future of autonomous systems
- For drones, this opens the door to move beyond simple GPS routes to true situational awareness
- Simultaneously, advances in NLP let us "talk" to AI in new ways

The Goal: To merge these fields, enabling a human to effortlessly collaborate with a drone

Methodology

Stage 1: Flight Log Generation

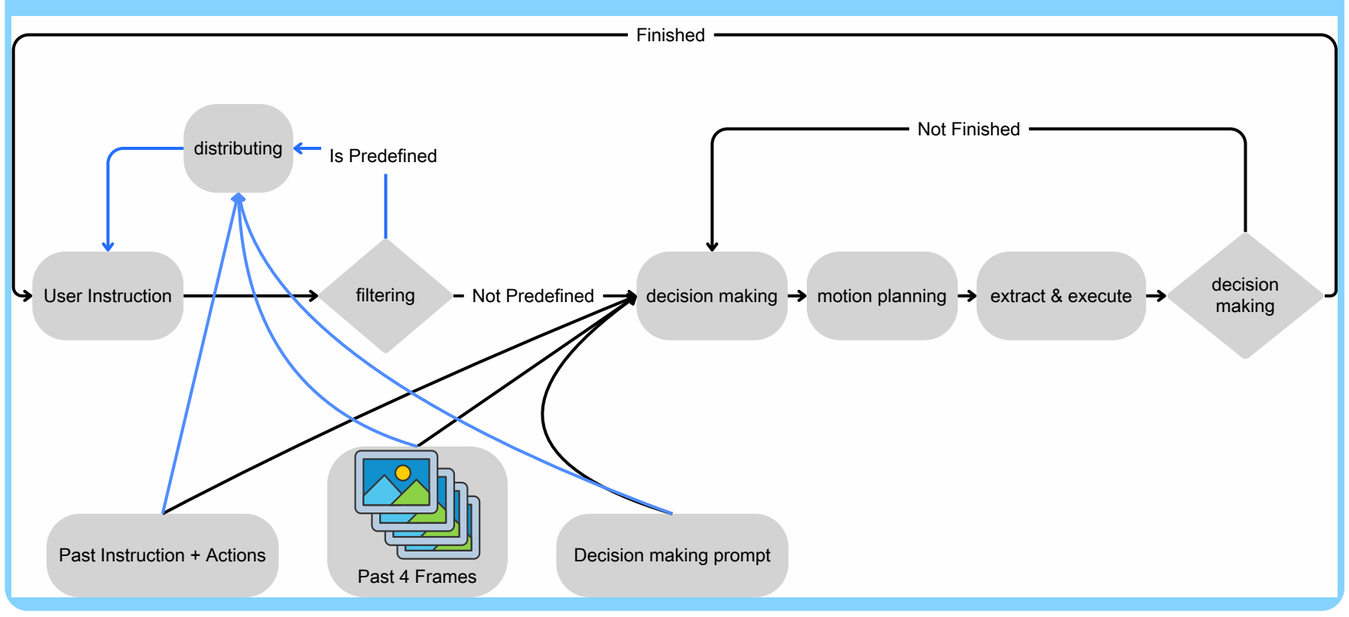
1. Test and select VLMs based on their ability to summarize a few consecutive frames
2. Determine the Flight Log Formation
3. Design 2 pipelines to complement each other
4. Merge (summarize) the 2 Flight Logs into the final one



Stage 2: Autonomous Flight

1. Setup AirSim Simulator (Unreal Engine + ROS2)
2. Customize drone API
3. Define instruction levels
4. Design a pipeline and sub-methods to convert high/middle level instructions to low level instructions

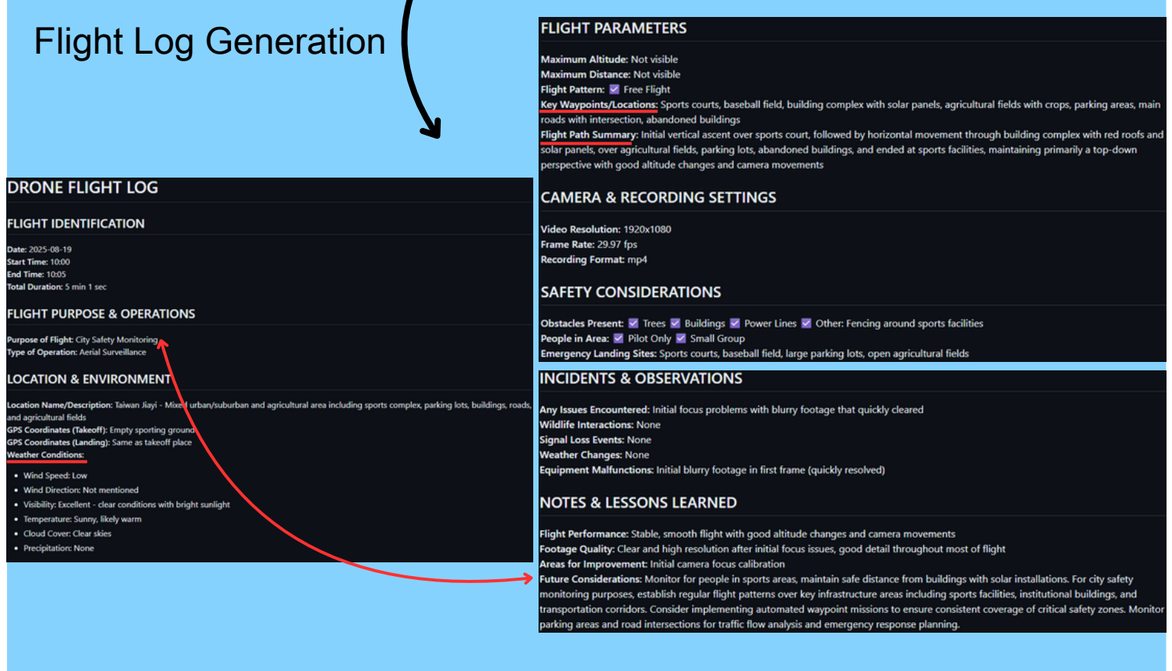
	Name	Description	Example
High Level	(Vision) Task	Ambiguous goals	Find me an old temple
High Level	Predefined Mission	Complicated mission	Return Flight
Middle Level	Navigation	Specific fly ideas	fly in a 100m square
Low Level	Action	Drone api	move_forward(x)



Results

The following result will be demonstrated in pictures. For a better experience, visit the GitHub page to check out the videos and logs

Flight Log Generation



DRONE FLIGHT LOG

FLIGHT IDENTIFICATION

Date: 2025-08-19
Start Time: 10:00
End Time: 10:05
Total Duration: 5 min 1 sec

FLIGHT PURPOSE & OPERATIONS

Purpose of Flight: City Safety Monitoring
Type of Operation: Aerial Surveillance

LOCATION & ENVIRONMENT

Location Name/Description: Taiwan Jiayi - Mixed urban/suburban and agricultural area including sports complex, parking lots, buildings, roads, and agricultural fields
GPS Coordinates (Takeoff): Empty sporting ground
GPS Coordinates (Landing): Same as takeoff place
Weather Conditions:

- Wind Speed: Low
- Wind Direction: Not mentioned
- Visibility: Excellent - clear conditions with bright sunlight
- Temperature: Sunny, likely warm
- Cloud Cover: Clear skies
- Precipitation: None

FLIGHT PARAMETERS

Maximum Altitude: Not visible
Maximum Distance: Not visible
Flight Pattern: ☑ Free Flight
Key Waypoints/Locations: Sports courts, baseball field, building complex with solar panels, agricultural fields with crops, parking areas, main roads with intersection, abandoned buildings
Flight Path Summary: Initial vertical ascent over sports court, followed by horizontal movement through building complex with red roofs and solar panels, over agricultural fields, parking lots, abandoned buildings, and ended at sports facilities, maintaining primarily a top-down perspective with good altitude changes and camera movements

CAMERA & RECORDING SETTINGS

Video Resolution: 1920x1080
Frame Rate: 29.97 fps
Recording Format: mp4

SAFETY CONSIDERATIONS

Obstacles Present: ☑ Trees ☑ Buildings ☑ Power Lines ☑ Other: Fencing around sports facilities
People in Area: ☑ Pilot Only ☑ Small Group
Emergency Landing Sites: Sports courts, baseball field, large parking lots, open agricultural fields

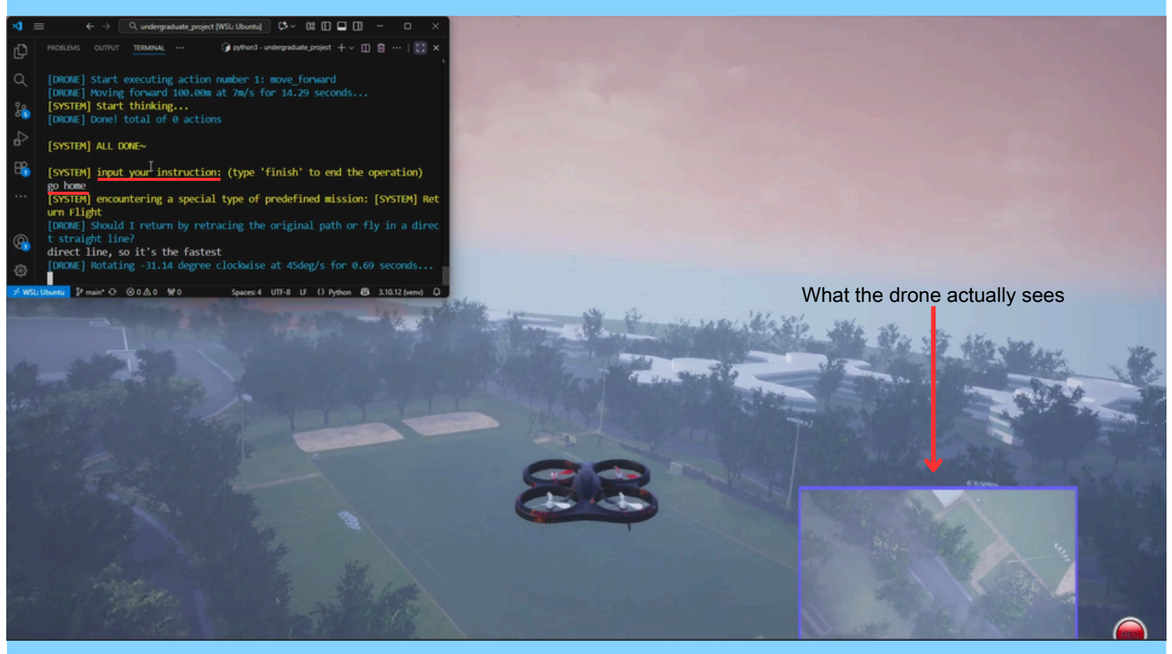
INCIDENTS & OBSERVATIONS

Any Issues Encountered: Initial focus problems with blurry footage that quickly cleared
Wildlife Interactions: None
Signal Loss Events: None
Weather Changes: None
Equipment Malfunctions: Initial blurry footage in first frame (quickly resolved)

NOTES & LESSONS LEARNED

Flight Performance: Stable, smooth flight with good altitude changes and camera movements
Footage Quality: Clear and high resolution after initial focus issues, good detail throughout most of flight
Areas for Improvement: Initial camera focus calibration
Future Considerations: Monitor for people in sports areas, maintain safe distance from buildings with solar installations. For city safety monitoring purposes, establish regular flight patterns over key infrastructure areas including sports facilities, institutional buildings, and transportation corridors. Consider implementing automated waypoint missions to ensure consistent coverage of critical safety zones. Monitor parking areas and road intersections for traffic flow analysis and emergency response planning.

Autonomous Flight



Conclusion

- VLMs establish a novel interface between humans and drones
- The Flight Log generation method provides a simple, effortless evaluation of drone footage
- The Autonomous Flight pipeline features a flexible architecture, allowing for the integration of sub-methods without disrupting the core framework
- Leveraging AirSim's integration of Unreal Engine and ROS2 for high-fidelity environmental and sensor simulation, this framework is transferable to real-world applications when the hardware requirements are met