



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

Author: 游宗謙, Advisor: 廖文宏博士
National Chengchi University, Taipei, Taiwan

Introduction

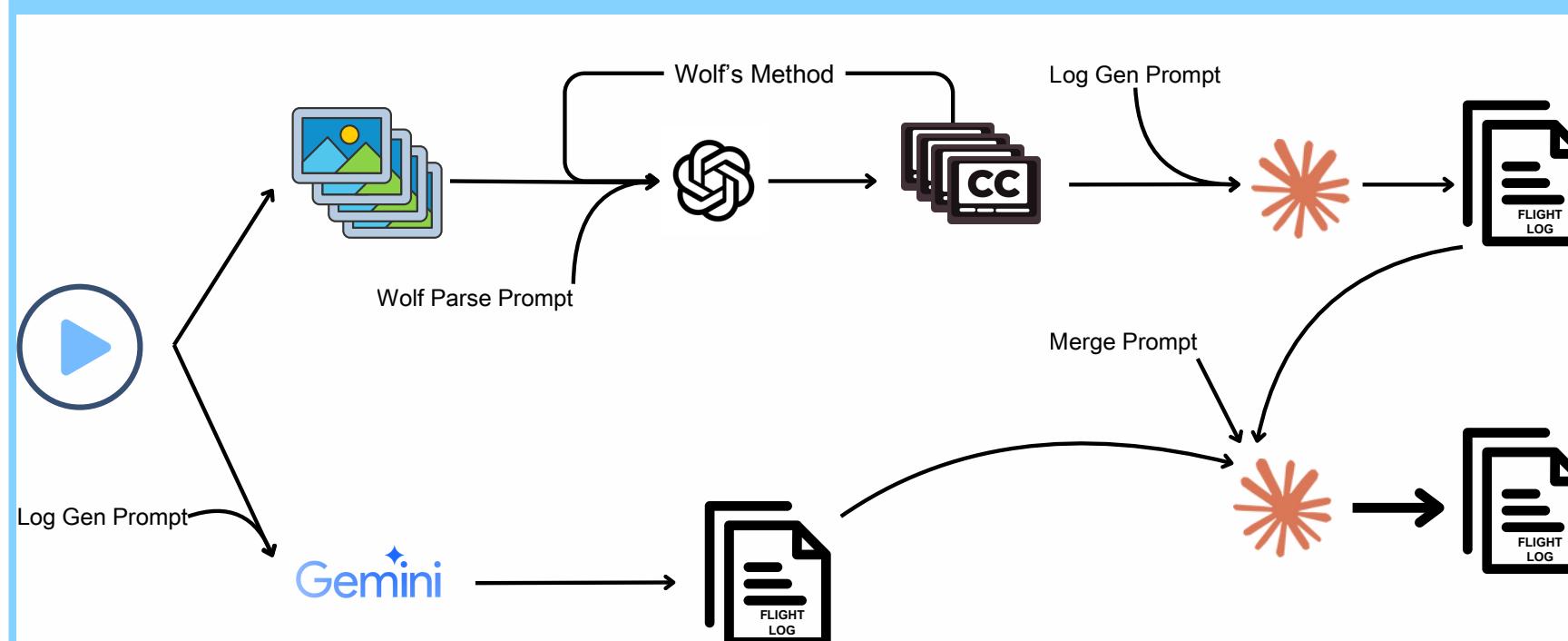
- 人工智能在電腦視覺領域的快速崛起，正大幅推動自主系統的未來發展
- 對無人機來說，這代表能從以往僅靠 GPS 的路線飛行，進化到真正具備情境理解能力
- 同時，自然語言處理的進步，讓我們能以全新方式「與 AI 對話」

目標: 結合這兩個領域，使人類能夠更輕鬆地與無人機協作

Methodology

Stage 1: Flight Log Generation

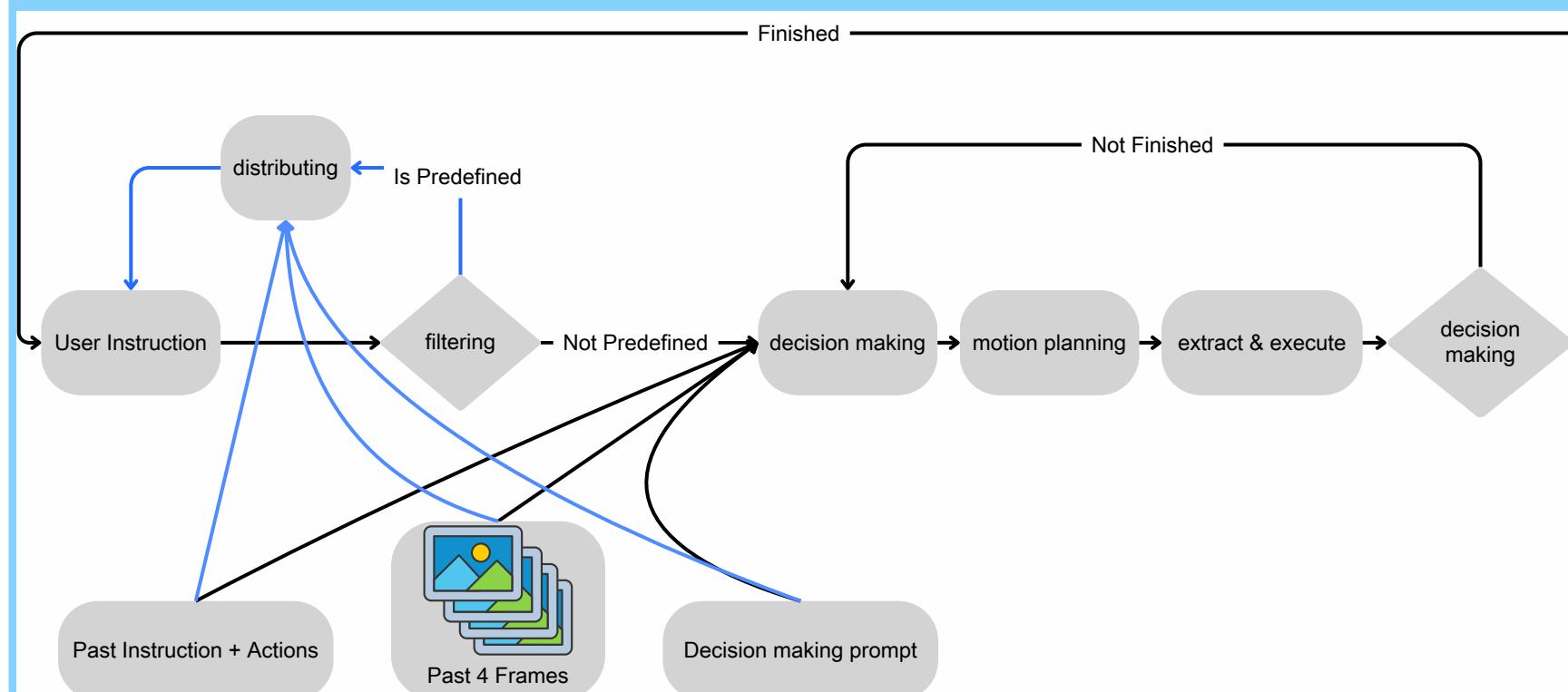
1. 基於各 VLM 總結連續畫面的能力，進行測試與選擇
2. 決定飛行日誌的格式
3. 設計兩條可互補的 pipeline
4. 將兩份飛行日誌合併成最終日誌



Stage 2: Autonomous Flight

1. 建立 AirSim (Unreal Engine + ROS2) 模擬環境
2. 自訂無人機 API
3. 定義使用者指令的層級
4. 設計流程與子方法，用以將高 / 中階指令轉換成低階指令

	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)

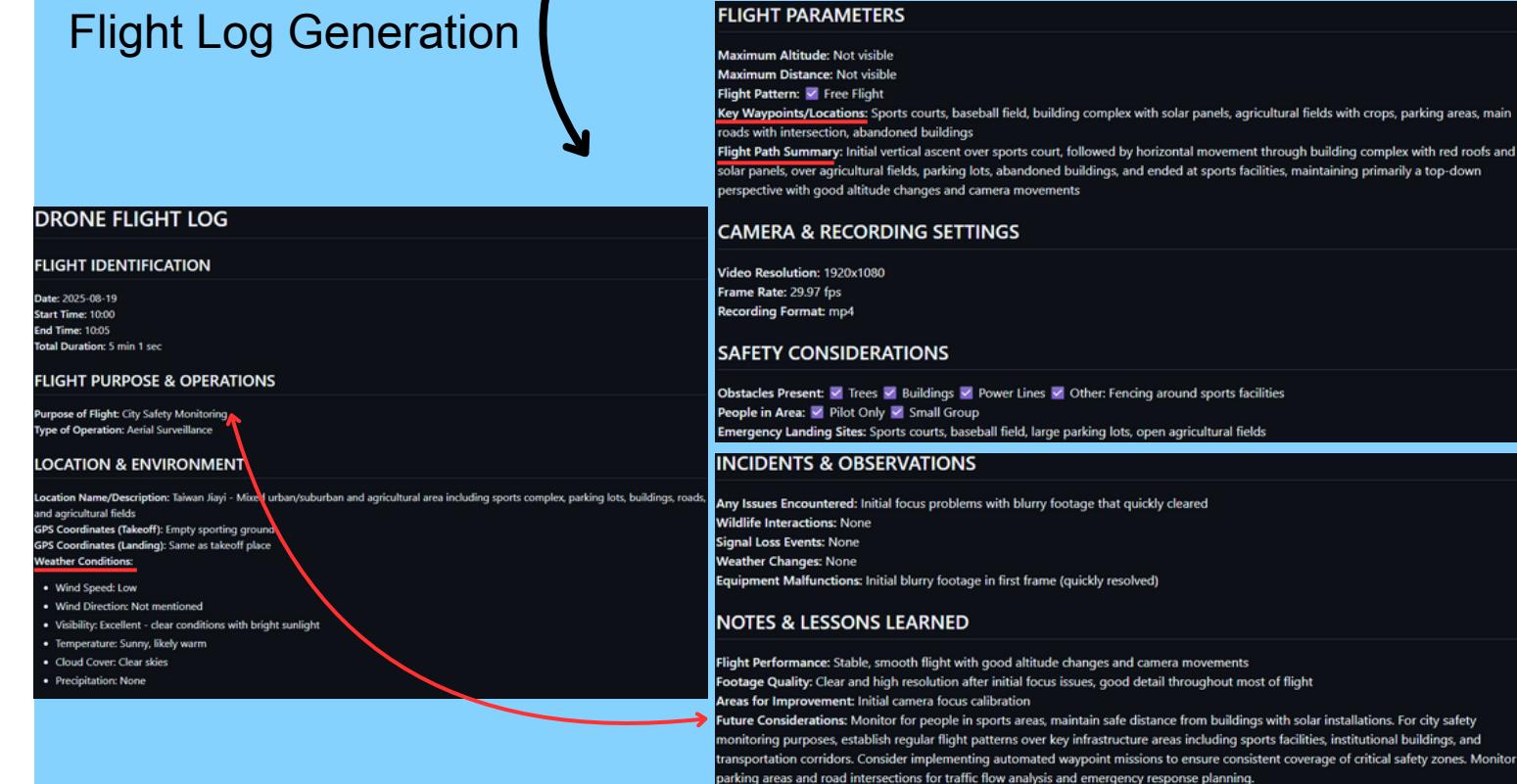


Results

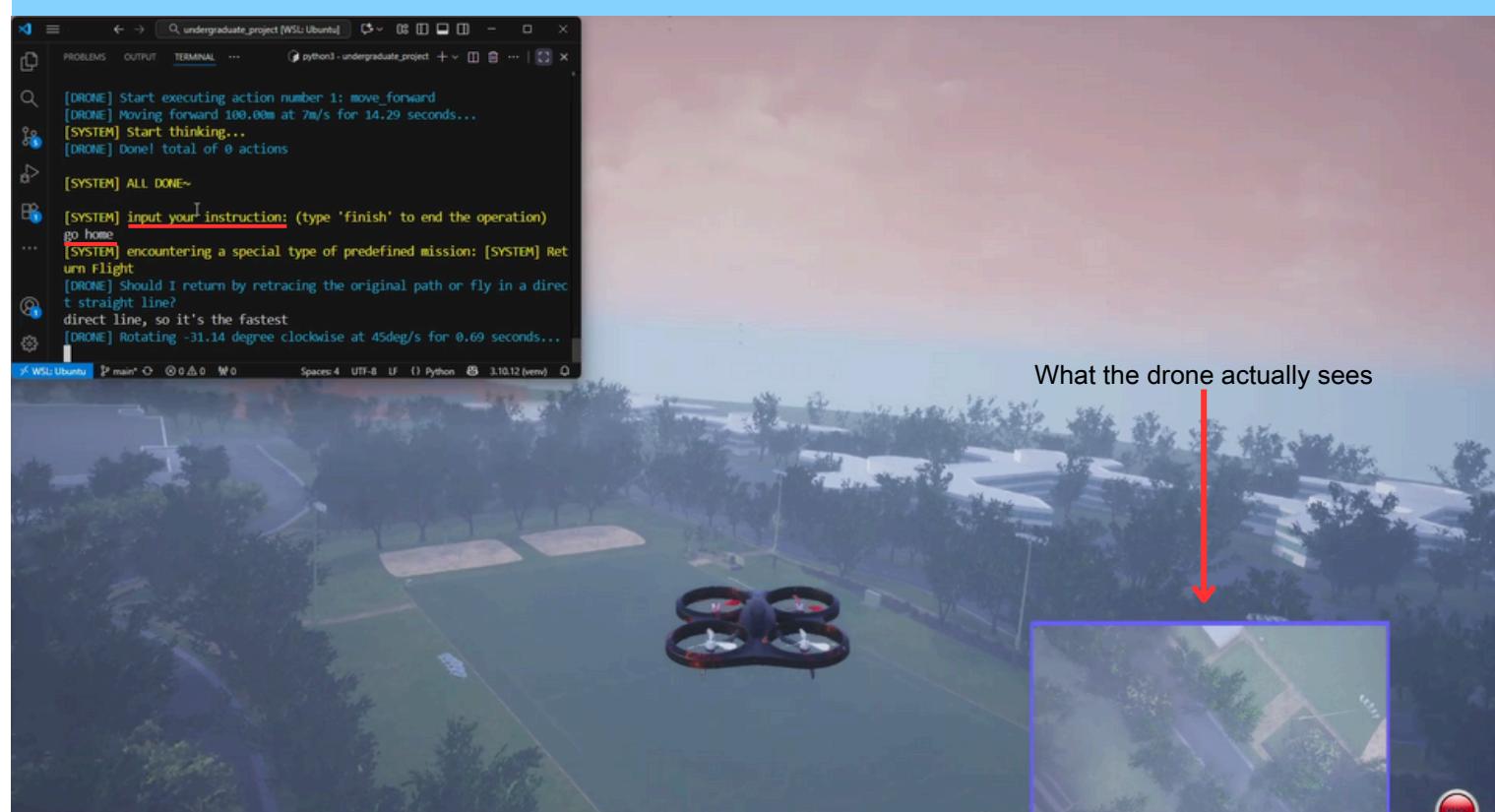
以下結果將以圖片展示，若想觀看完整的無人機影片與飛行日誌，可以到我的 [GitHub](#) 頁面進行瀏覽



Flight Log Generation



Autonomous Flight



Conclusion

- VLM 在人類與無人機之間建立了一種嶄新的介面
- 飛行日誌提供一種簡單、人力負擔低的無人機影像評估方式
- 自主飛行的 pipeline 具備高度彈性，可在不更動核心架構的前提下輕鬆擴充其他子方法
- 由於 AirSim 整合了 Unreal Engine 和 ROS2，其模擬環境與無人機輸入訊號皆高度擬真，一旦硬體需求滿足時，此架構即可直接接軌真實世界的應用