Day Final - Team 8 Ndvignition

From NTHU HSNL lab

OpenACC
More Science, Less Programming

Mentor

Rick Wu
From NVIDIA

Team Members

From NTHU HSNL lab

Ping-Chen Hsieh

HAN-YIN CHANG

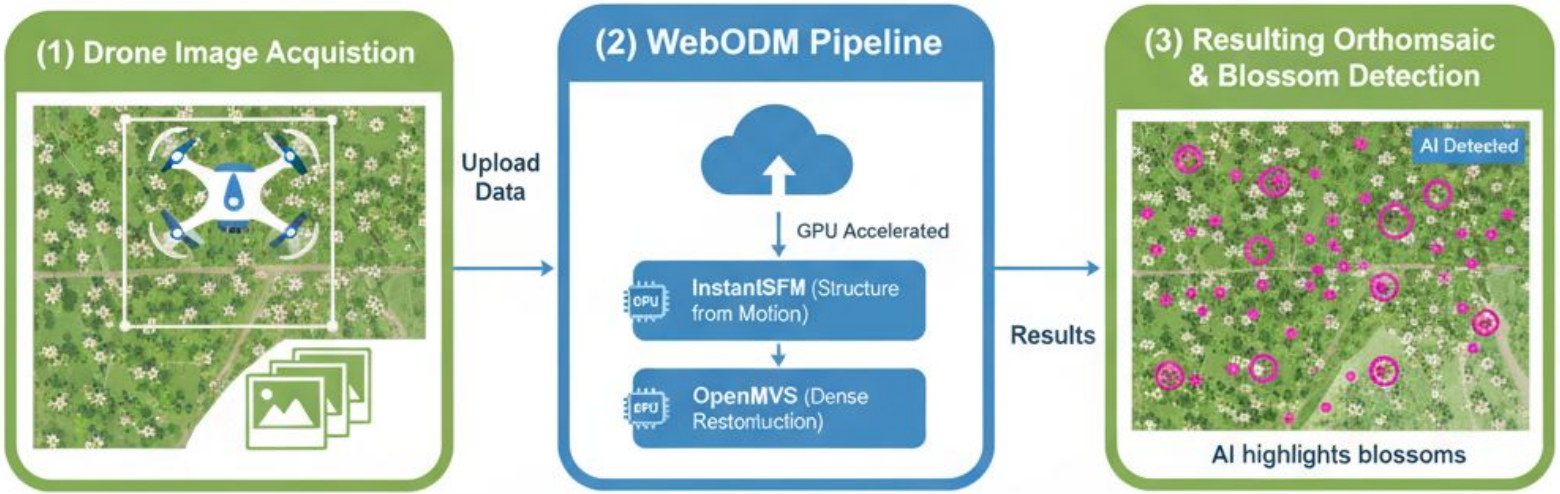
CHEN-YUAN CHUNG

Pin-Cheng Chan

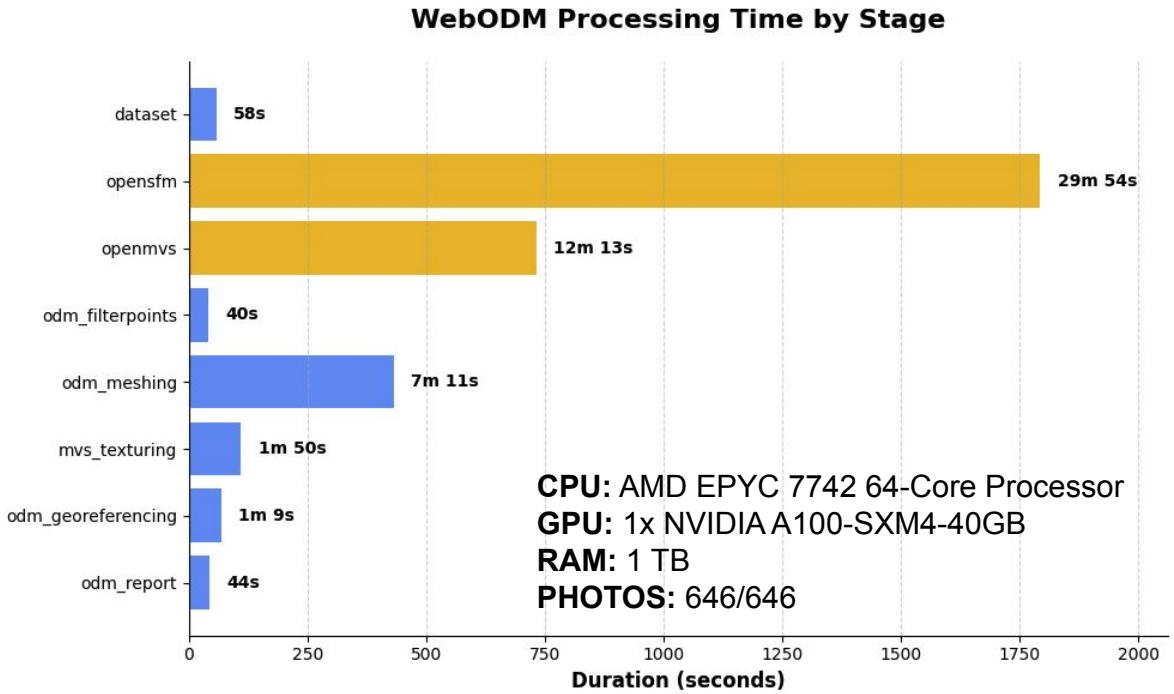
Ge-Han Wu

OpenACC More Science, Less Programming
OPEN HACKATHONS
NCHC 國家實驗研究院 國家高速網路與計算中心 National Center for High-performance Computing
NVIDIA

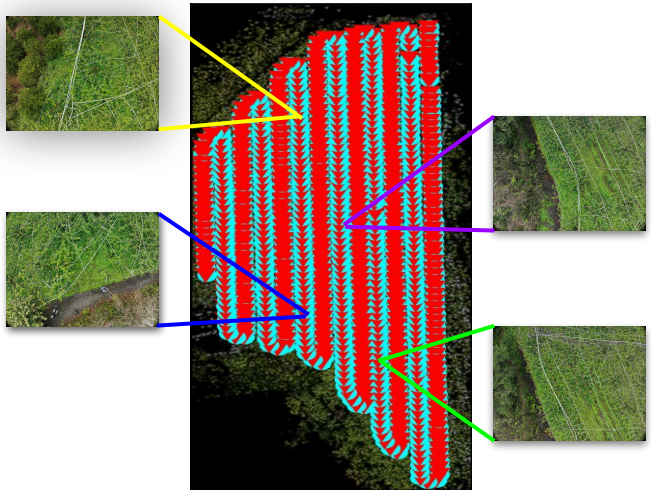
Acclerating WebODM Recorstuction with GPU (InstantSfM + OpenMVS)



Baseline Performance of WebODM

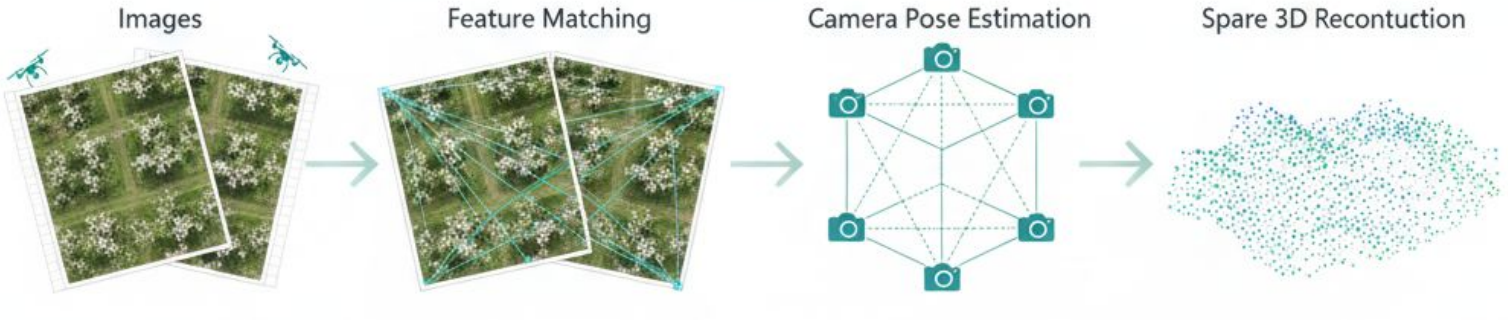


CPU: AMD EPYC 7742 64-Core Processor
GPU: 1x NVIDIA A100-SXM4-40GB
RAM: 1 TB
PHOTOS: 646/646

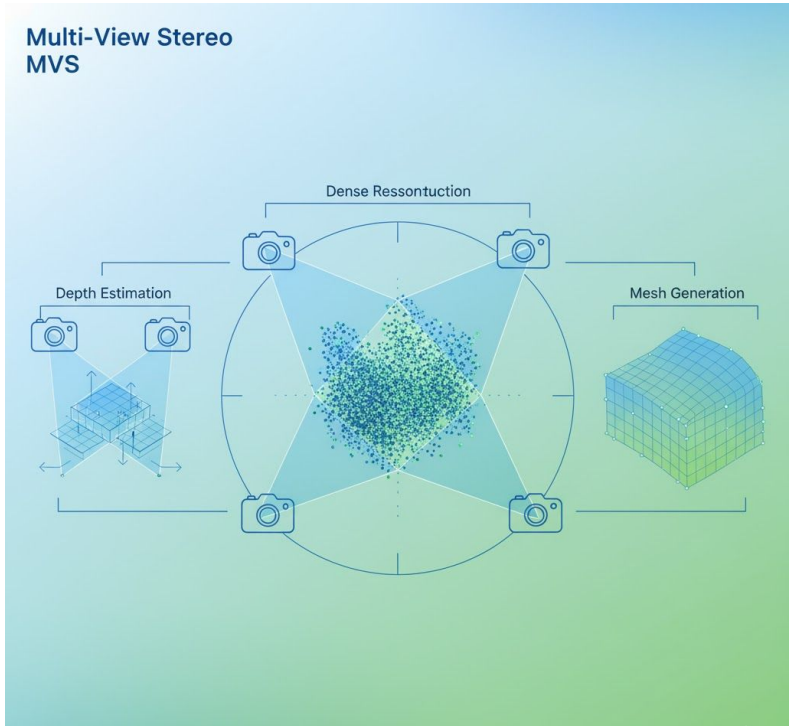


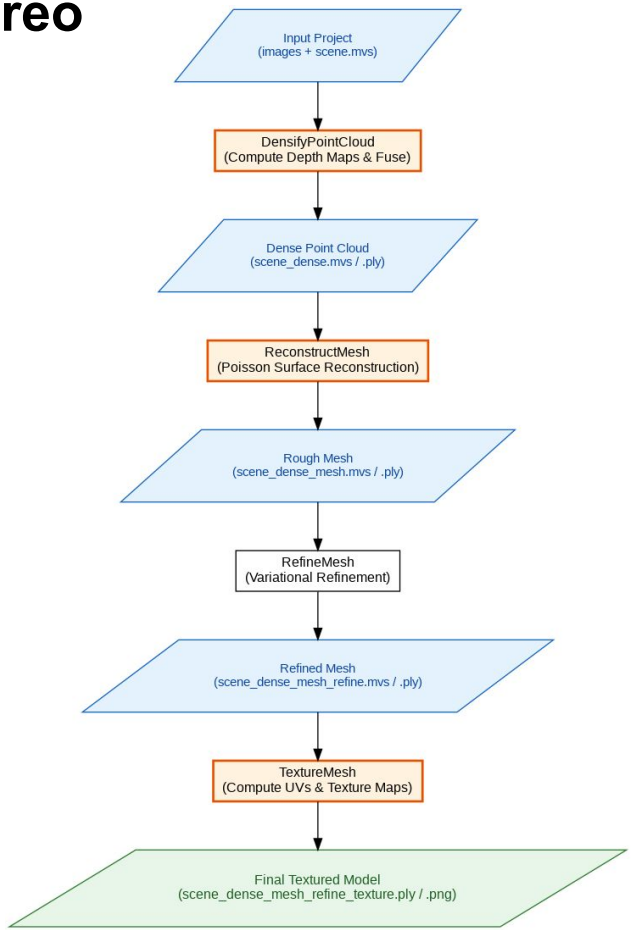
OpenACC More Science, Less Programming
OPEN HACKATHONS
NCHC 國家實驗研究院 國家高速網路與計算中心 National Center for High-performance Computing
NVIDIA

Structure from motion(SFM)

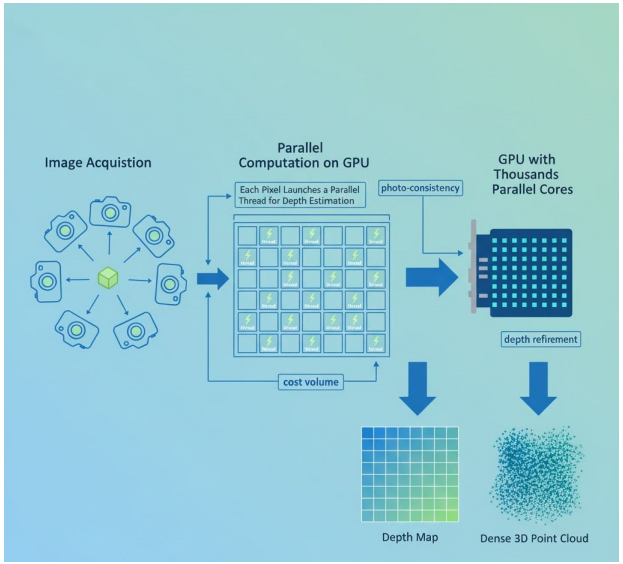


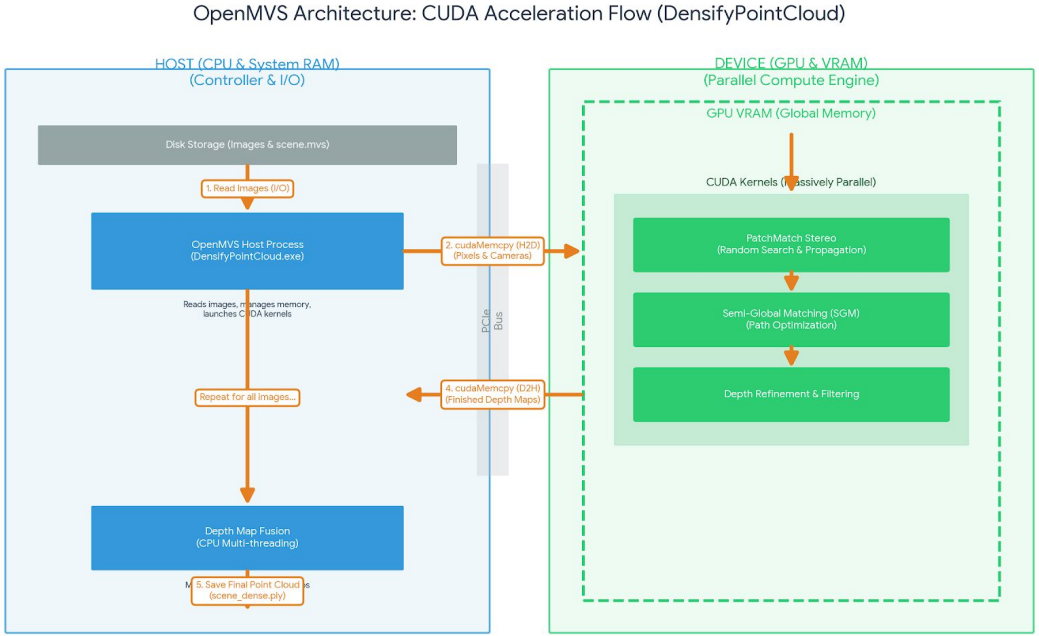
Multi-View Stereo



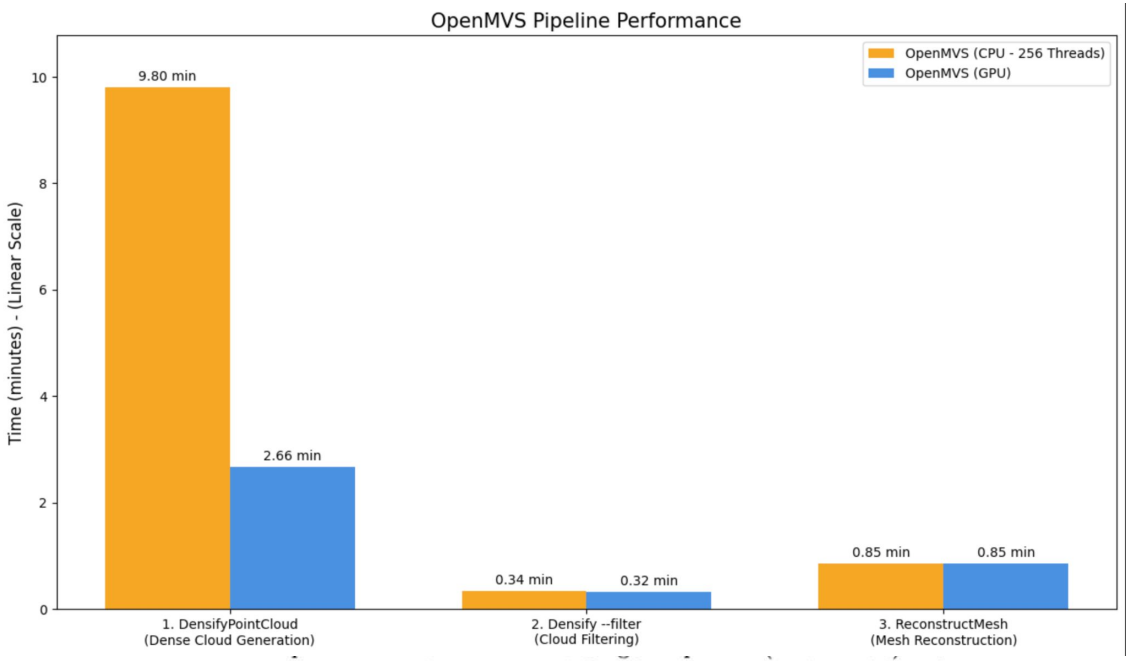


Why GPU Acceleration Works in MVS



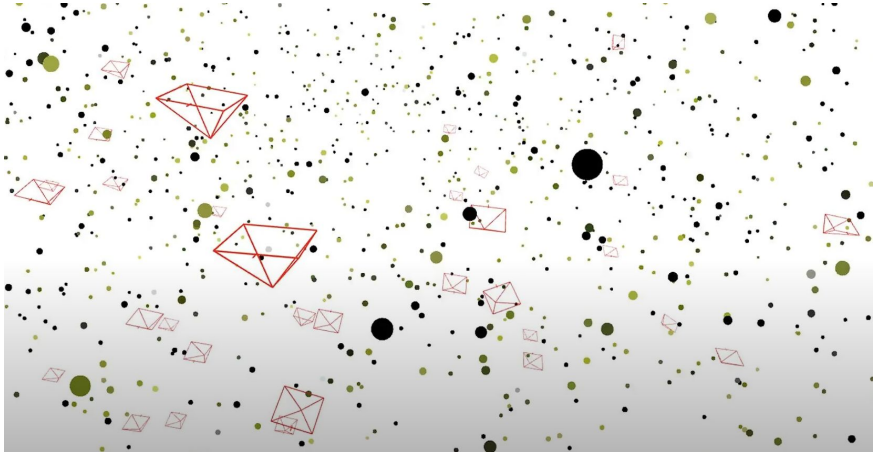


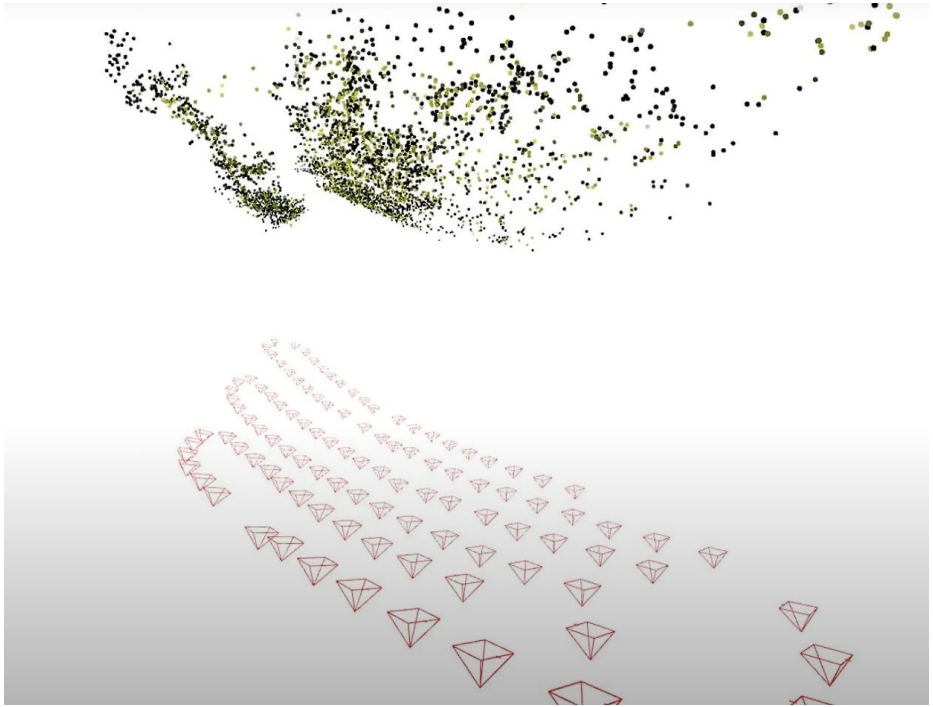
Openmvs Acceleration Results



Stage	OpenMVS (CPU - 256t) [min]	OpenMVS (GPU) [min]	Speedup
1. DensifyPointCloud	9.80	2.66	3.7×
2. Densify –filter	0.34	0.32	1.1×
3. ReconstructMesh	0.85	0.85	1.0×
4. Total Time	10.99	3.83	2.9×

Visual SFM

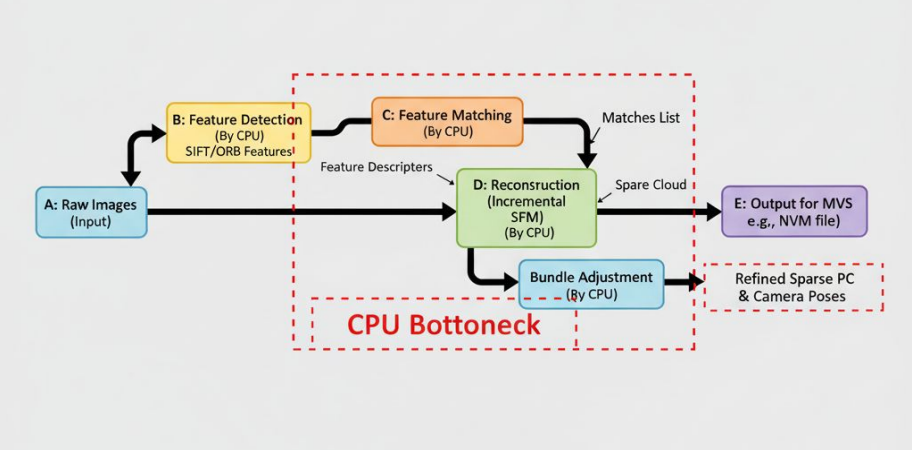


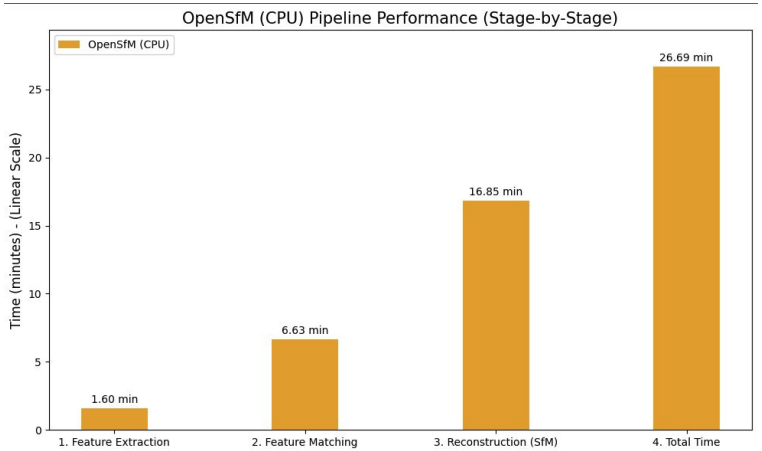


SfM Performance Bottleneck Analysis

OpenSFM

While incremental reconstruction benefits from the robustness of Ceres Solver, the preceding stages of feature extraction and matching still exhibit relatively slow performance.



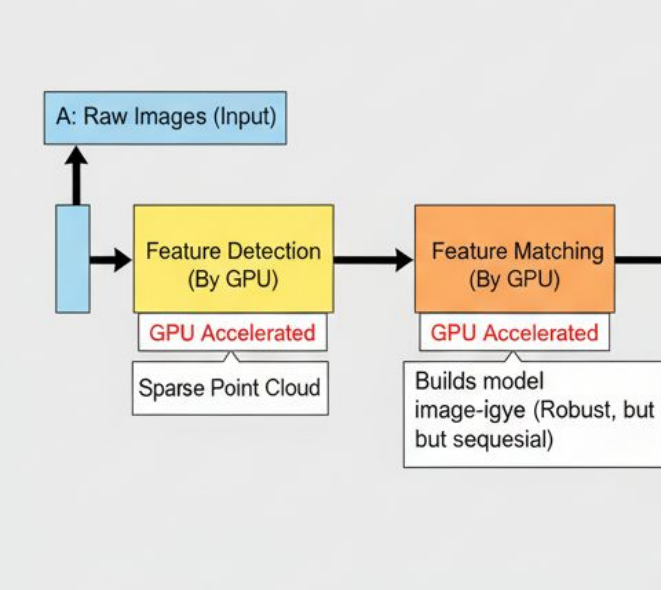


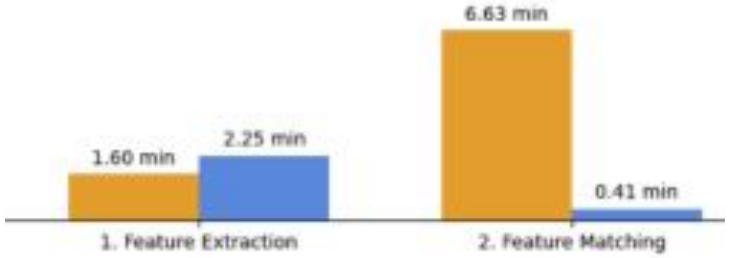
SfM Performance Bottleneck Analysis

Colmap

GPU Acceleration:

COLMAP can leverage the GPU to boost performance in feature detection & matching stage. “It speeds up matching by letting the GPU compare thousands of features at once instead of the CPU doing them one by one.”



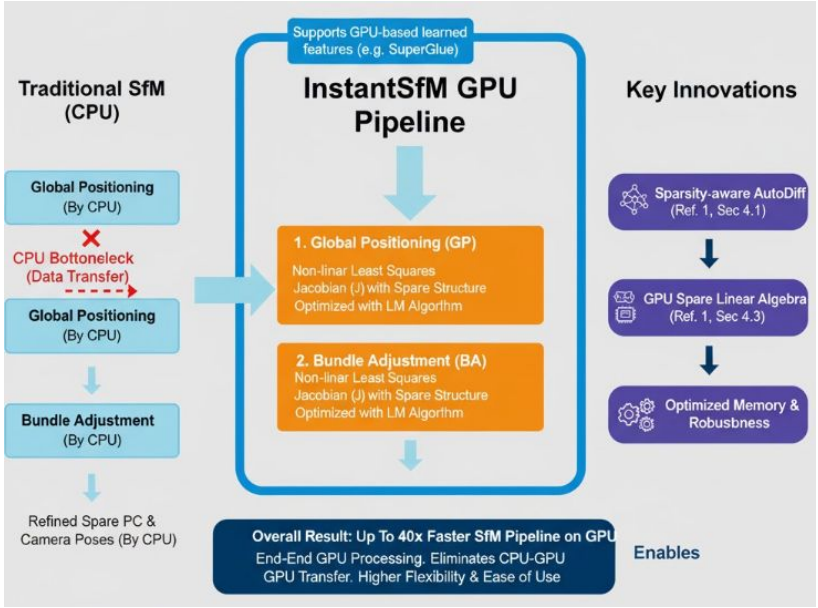


The Acceleration Evolution of SfM Pipelines

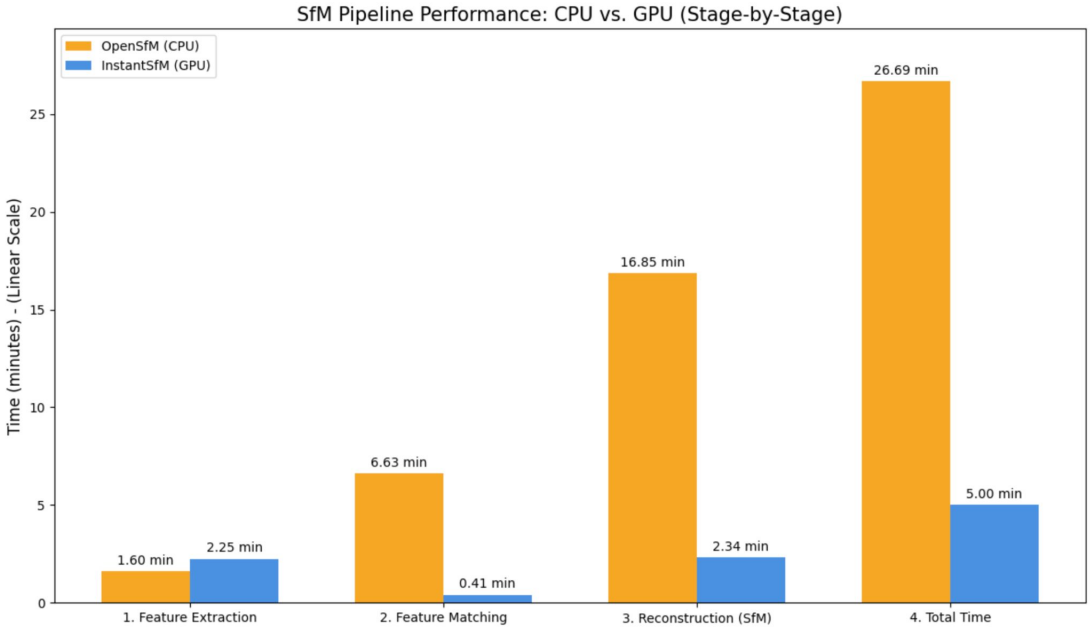
$$\boldsymbol{\theta} = \arg\min_{\mathbf{X},\boldsymbol{\zeta},\mathbf{K}} \sum_{i=1}^{C} \sum_{j=1}^{P} \rho(|| \underbrace{\Pi(\boldsymbol{\zeta}_i, \mathbf{X}_j, \mathbf{K}_i) - \mathbf{x}_{ij}}_{\mathbf{r}_{ij}} ||_2^2)$$

InstantSFM

- **Full GPU Execution:** InstantSfM fully shifts all computationally intensive non-linear optimization stages (GP & BA) of the SfM pipeline to the GPU within PyTorch Eager Mode.
- **Sparsity-Aware Acceleration:** It develops specialized GPU sparse auto-differentiation and linear solvers for the large and sparse Jacobian matrices in GP and BA, achieving significant speedups.



SFM Acceleration Results



Stage	OpenSfM (CPU) [min]	InstantSfM (GPU) [min]	Speedup
1. Feature Extraction	1.60	2.25	0.7× (slower)
2. Feature Matching	6.63	0.41	16.2×
3. Reconstruction (SfM)	16.85	2.34	7.2×
4. Total Time	26.69	5.00	5.3×

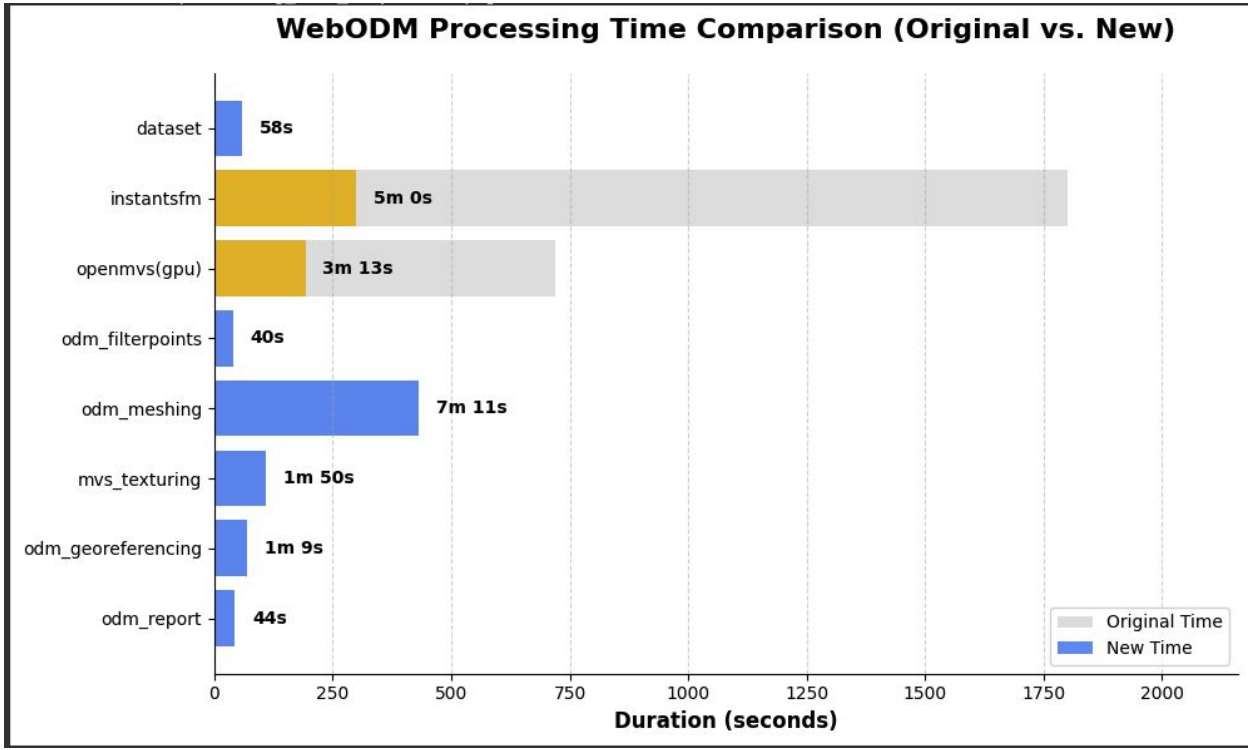
nsys profile

Feature Extraction

Feature Matching

Reconstruction

Overall speed up 2.6x



Energy Efficiency

INPUTS	
# CPU Cores	128
# GPUs (A100)	1
Application Speedup	5.3x

Node Replacement	42.7x

GPU NODE POWER SAVINGS			
	AMD Dual Rome 7742	**8x A100 80GB SXM4**	**Power Savings**
Compute Power (W)	46,992	6,500	40,492
Networking Power (W)	1,984	93	1,891
Total Power (W)	**48,976**	**6,593**	42,383

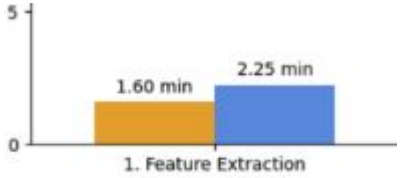
Node Power efficiency	**7.4x**

ANNUAL ENERGY SAVINGS PER GPU NODE			
	AMD Dual Rome 7742	**8x A100 80GB SXM4**	**Power Savings**
Compute Power (kWh/year)	411,650	56,940	354,710
Networking Power (kWh/year)	17,378	814	16,564
Total Power (kWh/year)	**429,028**	**57,754**	**371,274**

$/kWh	$ 0.18
Annual Cost Savings	$ 66,829.36
3-year Cost Savings	$ 200,488.09

Metric Tons of CO2	263
Gasoline Cars Driven for 1 year	57
Seedlings Trees grown for 10 years	4,351

Future Work: Accelerating the 3D Mesh Pipeline



current bottleneck

- Traditional algorithms are CPU and RAM-bound, struggling to process the massive (billions) dense point cloud.

- Requires re-reading all source images (high I/O) and performing complex view selection and color blending (high CPU) for millions of mesh faces.

Future

Distributed and Multi-GPU Reconstruction

Precision Agriculture Integration

Mobile Control Interface

What We Learned and How We Grew

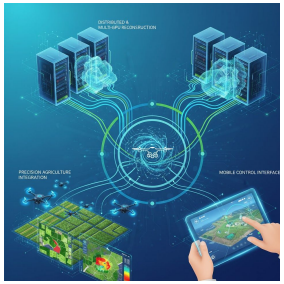
Identified key bottlenecks in WebODM's pipeline using log tracing and Nsight profiling, revealing that most time was spent in OpenSfM's CPU-based SfM stage.

Integrated InstantSfM to enable GPU acceleration, bridging data formats and achieving significant performance improvement.

Containerized the entire workflow with Docker for consistent, reproducible, and stable GPU execution across environments.

(Required) Create a storyline for publication on NCHC's website.

讓無人機影像拼接從小時變分鐘—GPU 加速農業航拍重建流程



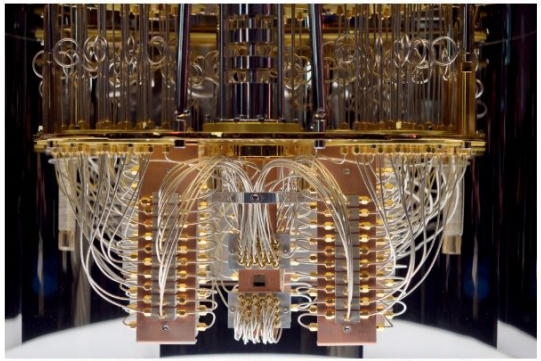
Ndvignition來自 黃能富老師帶領的高速網路實驗室，將無人機影像拼接加速了 5.36 倍！！

隨著無人機在農業與環境監測的應用越來越廣，影像拼接成為重要技術。傳統 WebODM 的重建流程在「結構重建（Structure-from-Motion）」與「密集重建（MVS）」階段耗時過長，難以滿足即時分析需求。本團隊針對效能瓶頸分析後，發現主要計算量集中於 OpenSfM。我們導入 GPU 平行化的 InstantSfM，它使用 PyTorch 架構進行 GPU 平行運算。InstantSfM 會把相機位置與特徵點之間的關係轉換成「稀疏矩陣」，只計算真正有關聯的部分，並利用 GPU 同時處理成千上萬個像素與相機參數，大幅加快求解速度。在 NVIDIA A100 上測試後，整體處理時間從近一小時縮短至約 20 分鐘。這項成果能讓精準農業、環境調查與城市數位孿生等領域的影像分析更即時、更有效率。

Stage	OpenSfM (CPU) [min]	InstantSfM (GPU) [min]	Speedup
1. Feature Extraction	1.60	2.25	0.7× (slower)
2. Feature Matching	6.63	0.41	16.2×
3. Reconstruction (SfM)	16.85	2.34	7.2×
4. Total Time	26.69	5.00	5.3×

報告投影片連結(由國網上傳到github)

量子算法模擬



haofan2023團隊成員來自臺灣大學資工系「洪士灝老師實驗室」，將量子演算法QAOA加速468倍！

— NVIDIA Mentors: Tian Zheng, Frank Lin, Yun-Yuan Wang

量子技術正以驚人的速度發展，預示著我們即將進入量子計算的時代。在這個過程中，量子電路模擬成為一個關鍵工具，它在量子硬體和軟體的開發中扮演著重要的角色，特別是在處理量子程式的編寫和驗證方面。傳統電腦的強模擬能夠獲得完整的量子狀態信息。這使得傳統電腦在構建量子系統方面變得不可或缺，尤其是在當前噪聲較多的中等規模量子（NISQ）時代。

量子近似優化算法（QAOA）是一種常用的量子算法，用於通過近似解來解決組合優化問題。然而，在虛擬量子計算機上執行QAOA對於解決需要大規模量子電路模擬的組合優化問題而言，會遇到模擬速度較慢的問題。團隊使用數學優化來壓縮量子操作，並結合有效的位元操作進一步降低計算複雜性，透過GPU加速最高獲取468倍的加速效果！

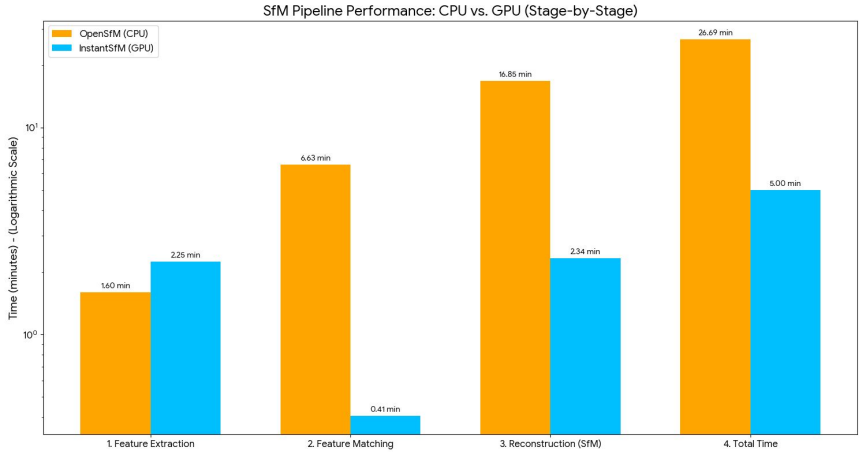
Table 1: The elapsed time of 5-level QAOA (unit: second, double).

Qubit	CPU_{Single}	$CPU_{Mutiple}$	CPU_{Cache}	GPU_{Cache}	GPU_{All}
23	29.80	1.28 (23x)	1.28 (63x)	0.24 (120x)	0.06 (**341x**)
24	68.00	3.46 (20x)	3.46 (43x)	0.55 (123x)	0.12 (**382x**)
25	152.52	15.32 (10x)	15.31 (45x)	1.19 (127x)	0.23 (**404x**)
26	330.69	33.83 (10x)	33.83 (56x)	2.60 (126x)	0.56 (**417x**)
27	712.26	72.66 (10x)	72.66 (54x)	5.59 (127x)	1.08 (**427x**)
28	1556.87	156.52 (10x)	156.52 (54x)	11.96 (130x)	2.17 (**445x**)
29	3325.55	335.09 (10x)	335.09 (49x)	25.73 (129x)	4.45 (**451x**)
30	7226.46	718.33 (10x)	718.33 (47x)	55.20 (130x)	9.22 (**468x**)

更多資訊請看：https://github.com/nqobu/nvidia/raw/main/20231207/Team02.pdf

Application Background

We focus on accelerating **drone image stitching** for agricultural and environmental monitoring using WebODM. The original pipeline is CPU-based and extremely time-consuming, especially in the Structure-from-Motion (SfM) stage.
Our goal is to integrate GPU-accelerated InstantSfM to enable near real-time reconstruction for UAV imagery.



Hackathon Objectives and Approach

Profiling / hot spot identification: Profiled WebODM pipeline, identified *OpenSfM* as the main bottleneck (≈30 mins per run).

Programming model: Replaced OpenSfM with **InstantSfM**, a GPU-parallelized SfM implementation.

Integration: Modified WebODM pipeline to support InstantSfM execution and verified correctness.

Performance tuning: Tested on NVIDIA A100 cluster for comparison with CPU baseline.

Technical Accomplishments and Impact

Integrated GPU-based InstantSfM into WebODM successfully.

Achieved significant reduction in reconstruction time (pending benchmark).

Identified path toward **multi-GPU and distributed** InstantSfM for future scalability.

Enables faster UAV mapping for **precision agriculture** and **environmental monitoring**, reducing hours of processing to minutes.


Thank You
OpenACC
More Science, Less Programming