|  |  |  |
| --- | --- | --- |
| |  |  | | --- | --- | |  | ĐẠI HỌC QUỐC GIA TP.HCM  TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN  KHOA CÔNG NGHỆ THÔNG TIN |   A picture containing logo  Description automatically generated  **BÁO CÁO ĐỒ ÁN**  **ĐỀ TÀI: KHẢO SÁT PHOTOMETRY (PHOTOGRAMMETRY)**  **VÀ ỨNG DỤNG ĐO ĐẠC TRONG NÔNG NGHIỆP**  **MÔN: XỬ LÝ ẢNH SỐ VÀ VIDEO SỐ**  Giáo viên: TS. Lý Quốc Ngọc  Nhóm: VDS\_22  Sinh viên thực hiện: 22120170 - Nguyễn Trung Kiên  22120295 - Hồ Minh Quang  22120307 - Lê Quang Vĩnh Quyền  22120311 - Lê Hoàng Sơn  22120437 - Lê Đình Hoàng Vũ  **TP. Hồ Chí Minh – Năm 2025** |
|  |

Mục lục

[**Chương 1. Giới thiệu** 4](#_Toc187186398)

[**1.1. Ý nghĩa về khoa học của chủ đề** 4](#_Toc187186399)

[**1.2. Ý nghĩa về ứng dụng của chủ đề** 4](#_Toc187186400)

[**1.3. Phát biểu bài toán [1][2]** 5](#_Toc187186401)

[**1.3.1. Input** 5](#_Toc187186402)

[**1.3.2. Framework** 5](#_Toc187186403)

[**1.3.3. Output** 6](#_Toc187186404)

[**1.4. Đóng góp** 6](#_Toc187186405)

[**1.4.1. Mục đích** 6](#_Toc187186406)

[**1.4.2. Cài đặt lại mô hình** 6](#_Toc187186407)

[**1.4.3. Thử nghiệm và so sánh** 7](#_Toc187186408)

[**1.4.4. Hướng dẫn triển khai** 7](#_Toc187186409)

[**Chương 2. Các công trình nghiên cứu liên quan** 7](#_Toc187186410)

[**2.1. A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] [3]** 7](#_Toc187186411)

[**2.2. Photogrammetric point cloud generation and surface interpolation for change detection [2016] [4]** 8](#_Toc187186412)

[**2.3. Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] [5]** 9](#_Toc187186413)

[**2.4. Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] [6]** 11](#_Toc187186414)

[**2.5. Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] [7]** 12](#_Toc187186415)

[**2.6. So sánh dựa trên tiêu chí** 13](#_Toc187186416)

[**Chương 3. Giải pháp [10]** 18](#_Toc187186417)

[**3.1.** **Trích xuất các đặc trưng bất biến** 19](#_Toc187186418)

[**3.2.** **So khớp ảnh** 19](#_Toc187186419)

[**3.3.** **So khớp đặc trưng** 20](#_Toc187186420)

[**3.4. Tái tạo cấu trúc (Structure from Motion)** 20](#_Toc187186421)

[**3.5. Ước lượng bản đồ độ sâu (Depth Map Estimation)** 21](#_Toc187186422)

[**3.6. Tạo khung lưới cho mô hình (Meshing)** 21](#_Toc187186423)

[**3.7. Khớp các chi tiết vào lưới (Texturing)** 22](#_Toc187186424)

[**Chương 4. Cài đặt và thử nghiệm chương trình** 22](#_Toc187186425)

[**Chương 5: Hạn chế và hướng phát triển** 24](#_Toc187186426)

[**5.1. Hạn chế** 24](#_Toc187186427)

[**5.2. Phương hướng phát triển** 24](#_Toc187186428)

[**Chương 6: Tham khảo** 25](#_Toc187186429)

# **Chương 1. Giới thiệu**

## **1.1. Ý nghĩa về khoa học của chủ đề**

* **Phạm vi nghiên cứu rộng lớn:**

Kết hợp cả vật lý và toán học để tạo ra, phân tích và diễn giải hình ảnh.

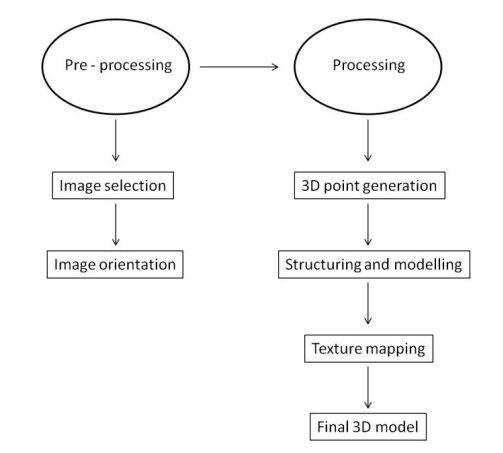
* **Nghiên cứu sự tương tác ánh sáng - vật thể:**  
  Photometry và Photogrammetry giúp hiểu sâu hơn về cách ánh sáng phản xạ, hấp thụ và tán xạ trên bề mặt vật thể. Điều này hỗ trợ việc xây dựng mô hình số liệu chính xác về môi trường tự nhiên.
* **Phát triển công nghệ hình ảnh và đo lường:**  
  Các kỹ thuật này thúc đẩy nghiên cứu các công nghệ xử lý ảnh, từ cảm biến quang học đến thuật toán tái tạo 3D, đóng góp cho nhiều lĩnh vực như viễn thám, đồ họa máy tính, và học máy.
* **Phân tích không gian và địa hình:**  
  Photogrammetry hỗ trợ phân tích và mô phỏng địa hình, cấu trúc cây trồng, và môi trường tự nhiên bằng cách tái tạo không gian ba chiều từ dữ liệu hình ảnh.
* **Nâng cao hiểu biết liên ngành:**  
  Kết nối giữa khoa học quang học, địa lý, toán học, và nông học mở ra nhiều hướng nghiên cứu đa lĩnh vực.

## **1.2. Ý nghĩa về ứng dụng của chủ đề**

* **Đo đạc diện tích và khối lượng cây trồng:**  
  Hỗ trợ xác định kích thước, mật độ, và sức khỏe cây trồng một cách tự động và chính xác, giúp lập kế hoạch sản xuất nông nghiệp hiệu quả hơn.
* **Theo dõi sức khỏe cây trồng:**  
  Sử dụng phân tích ánh sáng để đánh giá tình trạng dinh dưỡng và phát hiện sớm bệnh hoặc sự thay đổi bất thường trong môi trường canh tác.
* **Quản lý nguồn tài nguyên:**  
  Dữ liệu không gian giúp tối ưu hóa việc sử dụng nước, phân bón, và đất canh tác, giảm lãng phí và bảo vệ môi trường.
* **Hỗ trợ lập bản đồ và quy hoạch nông nghiệp:**  
  Tạo ra bản đồ chi tiết về địa hình và cây trồng phục vụ quy hoạch nông nghiệp chính xác, đặc biệt ở những khu vực khó tiếp cận.
* **Ứng dụng trong nông nghiệp thông minh:**  
  Kết hợp Photometry với trí tuệ nhân tạo và cảm biến IoT để tự động hóa các hoạt động như phun thuốc, thu hoạch, và quản lý mùa vụ.

## **1.3. Phát biểu bài toán [1][2]**

Photogrammetry là kỹ thuật tái tạo không gian 3D từ các bức ảnh 2D, ứng dụng rộng rãi trong nhiều lĩnh vực như nông nghiệp, xây dựng và quản lý tài nguyên. Quy trình này bao gồm 2 giai đoạn chính: tiền xử lý và xử lý. Mỗi giai đoạn đều đóng vai trò quan trọng trong việc đảm bảo độ chính xác và chất lượng của mô hình.



### **1.3.1. Input**

Hình ảnh thô thu được từ UAV

### **1.3.2. Framework**

**Giai đoạn Tiền xử lý:**

* Kiểm tra chất lượng hình ảnh: Bước này rất quan trọng để đảm bảo chỉ những hình ảnh chất lượng cao mới được sử dụng để tái tạo 3D. Cần loại bỏ các hình ảnh bị mờ, thiếu sáng, méo hình, hoặc bị che khuất bởi mây, bóng râm.
* Định hướng hình ảnh: Bước này xác định vị trí và góc chụp của từng hình ảnh trong không gian, sử dụng thông tin từ các điểm GCPs (Ground Control Points). GCPs là các điểm được khảo sát bằng GPS hoặc các thiết bị đo đạc chính xác khác, giúp liên kết các hình ảnh với hệ tọa độ thực tế. Quá trình định hướng hình ảnh bao gồm việc tìm kiếm các điểm đồng nhất trên nhiều hình ảnh và tính toán các tham số định hướng (bao gồm vị trí, góc xoay, và tiêu cự của camera).

**Giai đoạn Xử lý:**

* Tạo đám mây điểm(3D point generation): Sau khi định hướng hình ảnh, bước tiếp theo là tạo ra một đám mây điểm 3D. Đám mây điểm này được tạo ra bằng cách xác định các điểm ảnh tương ứng trên nhiều hình ảnh và sau đó sử dụng kỹ thuật tam giác để tính toán vị trí 3D của các điểm đó. Mật độ điểm của đám mây điểm phụ thuộc vào số lượng hình ảnh, mức độ chồng chéo giữa các hình ảnh, và thuật toán kết hợp được sử dụng.
* Xây dựng mô hình bề mặt (structuring and modelling) : Từ đám mây điểm 3D, một mô hình bề mặt được tạo ra. Bước này liên quan đến việc chọn một thuật toán phù hợp để kết nối các điểm với nhau, loại bỏ nhiễu và tạo ra một bề mặt liên tục. Thường thì là Delaunay Triangulation và Ball-Pivoting Algorithm.

Kết quả của bước này là một mô hình đa giác, thường được gọi là mesh, thể hiện hình dạng tổng quát của đối tượng.

Mô hình mesh có thể được tối ưu hóa để giảm kích thước và loại bỏ các tam giác dư thừa mà không làm giảm độ phân giải và hình học của mô hình.

* Kết cấu bề mặt (Texture mapping): Để tạo ra một mô hình 3D trực quan và chân thực, cần áp dụng kết cấu từ các hình ảnh ban đầu lên mô hình bề mặt. Quá trình này liên quan đến việc ánh xạ màu sắc và chi tiết từ các hình ảnh lên các đa giác của mô hình 3D.

### **1.3.3. Output**

* Mô hình 3D: Mô hình này đại diện cho hình dạng và kết cấu của đối tượng được khảo sát. Mô hình 3D có thể được sử dụng cho nhiều mục đích khác nhau, bao gồm hiển thị, phân tích, và in 3D.
* Ảnh ortho: Đây là loại ảnh chụp từ trên cao đã được chỉnh sửa hình học để loại bỏ méo hình và sai lệch do góc chụp và địa hình. Ảnh ortho cung cấp một cái nhìn chính xác về khu vực khảo sát và có thể được sử dụng cho nhiều mục đích khác nhau, bao gồm lập bản đồ, phân tích đất đai, và quản lý tài nguyên.

## **1.4. Đóng góp**

### **1.4.1. Mục đích**

Chúng tôi tự hào giới thiệu sự đóng góp quan trọng của chúng tôi đối với lĩnh vực "Photometry" thông qua việc cài đặt lại mô hình trên một công trình nghiên cứu đã có sẵn. Điều này không chỉ giúp chúng tôi hiểu rõ hơn về triển khai thực tế của các phương pháp hiện tại mà còn đưa ra những cái nhìn mới và cải tiến về "Photogrammetry" và ứng dụng của nó trong nông nghiệp.

### **1.4.2. Cài đặt lại mô hình**

Chúng tôi đã tiếp cận đề tài bằng cách sử dụng một công trình nghiên cứu đã được công bố. Qua đó, chúng tôi đã cải tiến và tối ưu hóa lại mô hình ban đầu để đảm bảo rằng nó phản ánh tốt nhất với nhu cầu cụ thể của đề tài của chúng tôi. Quá trình này không chỉ thách thức khả năng thí nghiệm của chúng tôi mà còn tạo ra những điểm độc đáo trong việc triển khai và thực hành.

### **1.4.3. Thử nghiệm và so sánh**

Chúng tôi đã tiến hành các thử nghiệm chi tiết để so sánh mô hình cài đặt lại của chúng tôi với công trình gốc và các phương pháp khác. Những kết quả này không chỉ cung cấp một cái nhìn sâu sắc về hiệu suất của chúng tôi mà còn chứng minh khả năng mở rộng và sửa lỗi của mô hình trong quá trình cài đặt lại.

### **1.4.4. Hướng dẫn triển khai**

Chúng tôi đã xây dựng một hướng dẫn chi tiết về cách triển khai mô hình "D" của chúng tôi trên nền tảng cụ thể đã sử dụng trong công trình nghiên cứu gốc. Hướng dẫn này không chỉ là một tài liệu kỹ thuật mà còn cung cấp các lời khuyên về tối ưu hóa và duy trì mô hình trong môi trường thực tế.

# **Chương 2. Các công trình nghiên cứu liên quan**

## **2.1. A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] [3]**

**Input**: Hình ảnh được thu thập từ một máy ảnh EOS-60D 18 MP CMOS Digital SLR với ống kính EF 50mm f/1.4, chụp từ nhiều góc độ khác nhau xung quanh cây táo đã được cắt tỉa và không có lá.

**Tiền xử lý**:

Kiểm tra chất lượng hình ảnh: chỉ những ảnh chất lượng cao mới được sử dụng

Định hướng hình ảnh:

* Xác định vị trí, góc chụp từng ảnh
* Tìm kiếm các điểm đồng nhất trên nhiều hình ảnh và tính toán các tham số định hướng (vị trí, góc xoay, tiêu cự camera) thông qua phần mềm PhotoScan Professional.

**Xử lý**:

Tạo đám mây điểm:

* Đám mây điểm được tạo ra bằng cách xác định các điểm ảnh tương ứng trên nhiều hình ảnh.
* Sử dụng kỹ thuật Structure-from-Motion (SfM) và Multi-View Stereo (MVS) để tính toán vị trí 3D của các điểm đó.

Xây dựng mô hình bề mặt:

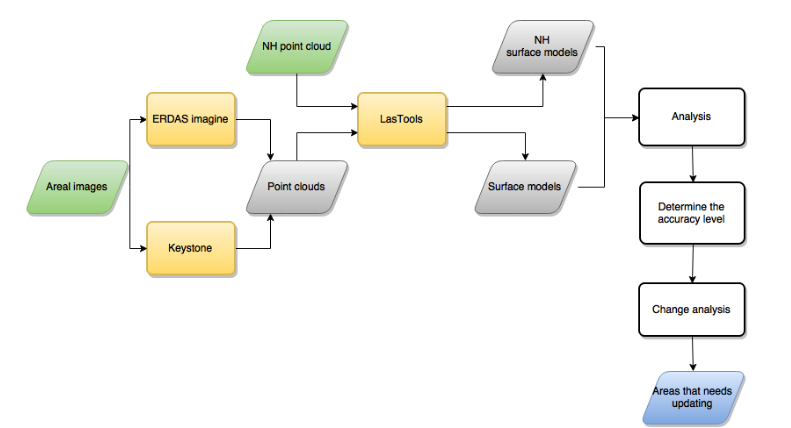
* Kết nối các điểm với nhau, loại bỏ nhiễu và tạo ra 1 bề mặt liên tục.
* Kết quả là một mô hình đa giác (mesh) thể hiện hình dạng tổng quát của cây.

Kết cấu bề mặt: Áp dụng kết cấu từ các hình ảnh ban đầu lên mô hình bề mặt để tạo ra một mô hình 3D hoàn chỉnh với thông tin màu sắc RGB.

**Output**:

Mô hình bề mặt (Digital Ground Truth - DGT) được tạo ra với độ chính xác cao, có thể được sử dụng để xác thực các hệ thống cảm biến khác trong việc mô tả thực vật 3D và hỗ trợ quyết định quản lý trong nông nghiệp chính xác.

## **2.2. Photogrammetric point cloud generation and surface interpolation for change detection [2016] [4]**



**Tiền xử lý**  
Chuẩn hóa dữ liệu ảnh và thông tin định hướng, loại bỏ các lỗi hệ thống, tối ưu hóa dữ liệu đầu vào cho quá trình ghép ảnh.

* Xử lý ảnh thô: Chuyển đổi dữ liệu ảnh thô từ camera sang định dạng .tif.
* Định hướng ảnh: Sử dụng phần mềm xử lý ảnh chuyên dụng (ví dụ ERDAS Imagine, Agisoft PhotoScan) để thực hiện định hướng ảnh.
* Định hướng tương đối: Xác định vị trí tương đối của các ảnh trong khối ảnh dựa trên các điểm tương đồng.
* Định hướng tuyệt đối: Sử dụng tọa độ của GCPs để gán hệ tọa độ và tỷ lệ cho khối ảnh.
* Tạo dữ liệu định hướng: Lưu thông tin định hướng của ảnh (tọa độ máy bay, độ cao, tham số biến đổi, v.v.) vào file .ori.

**Framework**

* Ghép ảnh (Image Matching): Xác định các điểm tương đồng (tie points) giữa các ảnh chồng chéo để xây dựng mô hình hình học của khu vực nghiên cứu.

+ Keystone chỉ được sử dụng để sản xuất một đám mây điểm từ hình ảnh chồng lấp. Để nhập hình ảnh và định hướng của chúng, một ứng dụng nhỏ độc lập đã được sử dụng.

+ ERDAS photogrammetry được sử dụng chủ yếu để tạo ra một đám mây điểm từ các hình ảnh chồng chéo. Hơn nữa, ERDAS Imagine được sử dụng để tạo ra các hình ảnh khác biệt và để xem kết quả cũng như thiết lập các phép chiếu phù hợp cho các bề mặt được tạo ra.

Kết quả thu được sau quá trình này các đám mây điểm 3D từ những điểm tương đồng trong không gian thu được từ giai đoạn 1.

Keystone sử dụng thuật toán Structure-from-Motion (SfM) sẽ tái tạo lại cấu trúc 3D của một cảnh từ nhiều ảnh 2D chồng chéo kết hợp với thuật toán SIFT (Scale-Invariant Feature Transform) để nhận dạng các keypoint, là các điểm đặc trưng trên ảnh không thay đổi theo tỷ lệ, góc xoay, hoặc độ sáng.

Sau khi xác định keypoint bằng SIFT, phương pháp RANSAC (Random Sample Consensus) được sử dụng để ghép các keypoint tương ứng giữa các ảnh và loại bỏ các outlier (các điểm không khớp).

* Tạo mô hình bề mặt (Creating Surface Model): LAStools được sử dụng để nội suy các đám mây điểm thành các bề mặt. Nội suy đám mây điểm 3D thành mô hình bề mặt liên tục, thể hiện địa hình khu vực nghiên cứu dựa trên bách phân vị 50, 90, 95 và 100. (50 là lấy trung vị, 90 95 100 là lấy giá trị cao hơn trong vùng lân cận, phù hợp với việc tạo mô hình bề mặt)

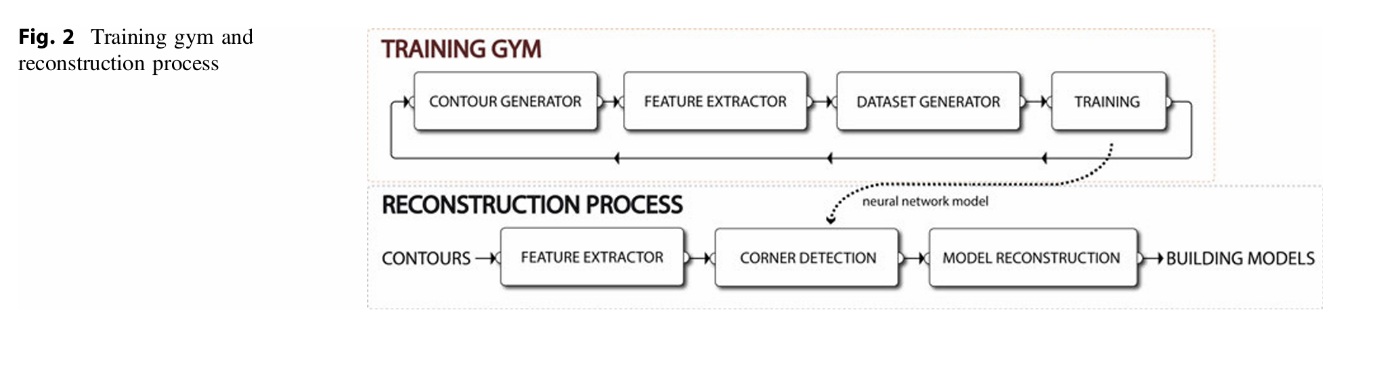
**Output**

Mô hình bề mặt (DSM) để cho giai đoạn phân tích nông nghiệp tiến hành.

## **2.3. Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] [5]**

Bài báo này trình bày một phương pháp tái tạo mô hình 3D từ dữ liệu trắc địa ảnh (photogrammetric 3D model) bằng cách sử dụng học sâu (deep learning). Mục tiêu là tạo ra các mô hình 3D đơn giản, hiệu quả, dễ sử dụng cho các kiến trúc sư và nhà thiết kế.

**Framework được chia thành 2 công đoạn chính:**



1. **Phòng tập (Training Gym):** Nơi tạo ra dữ liệu huấn luyện và huấn luyện mô hình.
2. **Quy trình tái tạo (Reconstruction Process):** Nơi sử dụng mô hình đã huấn luyện để tái tạo các mô hình 3D từ dữ liệu thực tế.

**2.3.1. Phòng tập (Training Gym):**

Mục tiêu: Tạo ra một mô hình mạng nơ-ron có khả năng phát hiện các góc trên đường viền 2D một cách chính xác.

Các công đoạn:

**2.3.1.1. Bộ tạo đường viền (Contour Generator):**

- Input: Không có input trực tiếp. Công cụ này hoạt động bằng cách sử dụng các tham số nội bộ và các seed ngẫu nhiên trong Grasshopper.

- Output: Các đường viền 2D khác nhau, với các biến thể về:

+ Hình dạng (ví dụ: hình chữ nhật, hình tròn, hình dạng tự do).

+ Độ cong của các cạnh.

+ Vị trí, góc độ và kích thước của các góc.

+ Nhiễu.

- Công đoạn xử lý:

+Tạo các đường viền dựa trên các hình dạng cơ bản và các biến đổi hình học.

+ Thêm nhiễu vào các đường viền để tăng tính đa dạng và khả năng chống chịu.

+ Chia đường viền thành 100 điểm nút.

**2.3.1.2. Bộ trích xuất đặc trưng (Feature Extractor):**

- Input: Các đường viền 2D đã được tạo ra ở bước 1.1 (dưới dạng tọa độ của các nút).

- Output: Một chuỗi các giá trị radian (góc) tương ứng với các nút trên đường viền.

- Công đoạn xử lý:

+ Tính toán góc ngoài tại mỗi nút trên đường viền.

+ Gán giá trị radian tương ứng cho mỗi nút (giá trị dương nếu quay theo chiều kim đồng hồ, âm nếu ngược chiều kim đồng hồ).

+ Chuẩn hóa dữ liệu (đưa giá trị về khoảng từ -π đến +π).

**2.3.1.3. Bộ tạo tập dữ liệu (Dataset Generator):**

- Input: Các đường viền 2D và chuỗi các giá trị radian tương ứng.

- Output: Các cặp dữ liệu (input, label) cho quá trình huấn luyện:

+ Input: Chuỗi các giá trị radian đã chuẩn hóa.

+ Label: 0 nếu là điểm góc, 1 nếu không phải điểm góc.

- Công đoạn xử lý:

+ Xác định các điểm góc thực tế trên đường viền.

+ Gắn nhãn (label) cho từng điểm trên đường viền, dựa trên việc so sánh với vị trí điểm góc.

+ Tạo các cặp input-label, đưa các giá trị radian và label về đúng thứ tự để huấn luyện mô hình.

**2.3.1.4. Đào tạo (Training):**

- Input: Các cặp dữ liệu (input, label) đã được tạo ở bước 1.3.

- Output: Mô hình mạng nơ-ron đã được huấn luyện (dưới dạng file pkl).

- Công đoạn xử lý:

+ Sử dụng các cặp input-label để huấn luyện mô hình mạng nơ-ron.

+ Cập nhật trọng số của mô hình sau mỗi epoch để giảm sai số và tăng độ chính xác.

+ Lưu lại mô hình đã được huấn luyện sau khi hoàn thành quá trình.

**2.3.2. Quy trình tái tạo (Reconstruction Process):**

Mục tiêu: Sử dụng mô hình đã huấn luyện để tái tạo mô hình 3D từ dữ liệu trắc địa ảnh.

Các công đoạn:

**2.3.2.1. Đầu vào (Input):**

- Input: Các đường viền 2D từ dữ liệu trắc địa ảnh thực tế (dưới dạng tọa độ của các nút).

**2.3.2.2. Bộ trích xuất đặc trưng (Feature Extractor):**

- Input: Các đường viền 2D từ dữ liệu thực tế.

- Output: Chuỗi các giá trị radian tương ứng với các nút trên đường viền.

- Công đoạn xử lý: Tương tự như bước 1.2, nhưng áp dụng cho dữ liệu trắc địa ảnh.

**2.3.2.3. Phát hiện góc (Corner Detection):**

- Input: Chuỗi các giá trị radian từ bước 2.2.

- Output: Các vị trí của các góc trên đường viền (các điểm có khả năng cao là góc).

- Công đoạn xử lý:

+ Đưa chuỗi các radian vào mô hình mạng nơ-ron đã huấn luyện.

+ Mô hình sẽ dự đoán xem mỗi điểm có phải là góc hay không (dựa trên đầu ra của hàm sigmoid).

+ Áp dụng ngưỡng để xác định các điểm được coi là góc.

**2.3.2.4. Tái tạo mô hình (Model Reconstruction):**

- Input: Vị trí các góc và các đoạn đường viền giữa các góc.

- Output: Mô hình 3D đã được tái tạo và đơn giản hóa.

- Công đoạn xử lý:

+ Phân đoạn đường viền dựa trên các vị trí góc đã được phát hiện.

+ Sử dụng các điểm góc làm điểm điều khiển cho các đường cong Bézier.

+ Tạo các bề mặt từ các đường cong Bézier.

+ Kết nối các bề mặt để tạo thành mô hình 3D hoàn chỉnh.

+ Đơn giản hóa các bề mặt để giảm số lượng mặt.

**2.3.2.5. Đầu ra (Output):**

+ Output: Mô hình 3D đã được tái tạo và đơn giản hóa (có thể ở các định dạng khác nhau).

## **2.4. Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] [6]**

**Input:** Hình ảnh lập thể của các đoạn kênh được chụp bằng máy ảnh kỹ thuật số đã được hiệu chuẩn dưới dạng JPEG.

**Đặc điểm:** Mỗi vị trí giám sát được chụp ảnh từ hai góc nhìn: từ hạ lưu nhìn ngược dòng và từ thượng lưu nhìn xuống dòng, đảm bảo tất cả bốn GCP được nhìn thấy trong mỗi ảnh.

**Tiền xử lý**

**Xử lý Hình ảnh:** Sử dụng thuật toán Biến đổi Đặc trưng Bất biến Tỷ lệ (SIFT) để tự động khớp các điểm trong ảnh. Xác định thủ công các GCP trong mỗi ảnh và so sánh chúng với dữ liệu khảo sát.

**Xử lý**

* **Tạo Đám mây Điểm:** Tạo một đám mây điểm 3D dày đặc đại diện cho bề mặt kênh bằng cách sử dụng tùy chọn "Tạo Bề mặt Dày đặc" cho cặp ảnh lập thể
* **Đánh giá Chất lượng Dữ liệu:** So sánh các mô hình được nhân bản của cùng một bề mặt và so sánh độ cao của VP được khảo sát với các điểm gần nhất của chúng trong mỗi đám mây điểm.
* **Xác định Thay đổi Thể tích:**So sánh đám mây điểm theo thời gian bằng tập lệnh Python tùy chỉnh. Nội suy cả hai đám mây điểm thành DEM lưới 5 mm. Tính toán tổng thay đổi thể tích trong vị trí giám sát giữa các ngày chụp ảnh dựa trên chênh lệch độ cao của các ô lưới tương ứng.

**Đầu ra:** Tập hợp các thay đổi thể tích được tính toán cho mỗi vị trí giám sát theo thời gian.

**4. Output**

**Ước tính Xói mòn Khe rãnh Tạm thời:** Nội suy dựa trên Phương pháp Diện tích Cuối được sửa đổi. Chuyển đổi thay đổi thể tích thành thông lượng khối lượng bằng cách sử dụng mật độ khối lượng đất mặt.

**Đầu ra:** Tốc độ xói mòn khe rãnh tạm thời cho toàn bộ kênh theo thời gian, được biểu thị bằng thông lượng khối lượng trên một đơn vị diện tích.

## **2.5. Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] [7]**

**Input**:

Hình ảnh thô từ các nguồn khác nhau

* Có nhãn: Ảnh từ các bộ dữ liệu như BlendedMVS, MegaDepth, DIML, thu thập với nhãn độ sâu từ stereo hoặc SfM.
* Không nhãn: Ảnh từ các tập công khai lớn như BDD100K, SA-1B, Open Images.

**Tiền xử lý**:

Thu thập dữ liệu:

* Dữ liệu có nhãn (1.5 triệu ảnh): Bao gồm các tập BlendedMVS, MegaDepth, DIML.
* Dữ liệu không nhãn (62 triệu ảnh): Lấy từ các tập như BDD100K, SA-1B.

Tạo nhãn giả:

* Sử dụng mô hình giáo viên được huấn luyện trên dữ liệu có nhãn để gán nhãn giả cho ảnh không nhãn.
* Gắn nhãn giả cho 62 triệu ảnh từ các tập dữ liệu công khai.

Tăng cường dữ liệu:

* Thực hiện các nhiễu mạnh như CutMix (kết hợp vùng từ hai ảnh), biến dạng màu sắc, và làm mờ Gaussian để tăng khả năng bền vững của mô hình khi gặp điều kiện hình ảnh đa dạng.

**Xử lý:**

Huấn luyện giáo viên:

* Sử dụng dữ liệu có nhãn (1.5 triệu ảnh) để huấn luyện mô hình giáo viên với tổn thất bất biến tỷ lệ (Affine-invariant loss).
* Các tham số định hướng (vị trí, góc xoay) của ảnh được tối ưu hóa để đảm bảo tính chính xác.

Huấn luyện sinh viên

* Dữ liệu kết hợp: Cả ảnh có nhãn và ảnh nhãn giả với tỷ lệ 1:2.
* Mục tiêu học tập thách thức: Áp dụng các biến dạng mạnh để buộc mô hình học biểu diễn mạnh mẽ hơn.

Căn chỉnh đặc trưng (Feature Alignment):

* Duy trì thông tin ngữ nghĩa từ encoder tiền huấn luyện (DINOv2).
* Tổn thất căn chỉnh đặc trưng giúp cân bằng giữa thông tin độ sâu và ngữ nghĩa, tăng hiệu suất trên cả nhiệm vụ độ sâu và phân đoạn ngữ nghĩa.

**Output**: Bản đồ độ sâu

## **2.6. So sánh dựa trên tiêu chí**

**2.6.1. Nguyên lý:**

|  |  |
| --- | --- |
| **Công trình** | **Nguyên lý** |
| Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] | Sử dụng Trắc địa ảnh tầm gần để tạo các đám mây điểm ảnh hoặc mô hình độ cao kĩ thuật số(DEM).  Từ đó, dữ liệu sẽ được cập nhật sau mỗi khoảng thời gian nhất định và được so sánh với các DEM đã tạo trước đó để kiểm tra độ sai khác của địa hình. |
| Photogrammetric point cloud generation and surface interpolation for change detection [2016] | So sánh các điểm ảnh từ nhiều hình ảnh chụp cùng một khu vực từ các góc độ khác nhau.  Bằng cách xác định các điểm ảnh tương ứng (tie points) trên các hình ảnh, thuật toán có thể tính toán khoảng cách giữa máy ảnh và điểm ảnh.  Từ đó xây dựng một mô hình 3D của khu vực được chụp. |
| Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] | Phát hiện góc: Phương pháp này tập trung vào việc phát hiện các góc trên các đường viền của các đối tượng (chủ yếu là các tòa nhà).  Phân đoạn đa tuyến: Sau khi phát hiện góc, các đường viền được chia thành các phân đoạn đa tuyến.  Sử dụng góc làm đặc trưng: Thay vì sử dụng trực tiếp hình ảnh hoặc dữ liệu phức tạp, phương pháp này sử dụng chuỗi các góc ngoài của đường viền làm đặc trưng để huấn luyện mô hình học sâu.  Tái tạo bề mặt: Từ các đường viền và phân đoạn, các bề mặt 3D được tái tạo bằng cách sử dụng các đường cong Bézier.  Hệ thống phòng tập (training gym): Sử dụng công cụ thiết kế tạo sinh Grasshopper để tạo ra các tập dữ liệu huấn luyện đa dạng và tự động, không cần dữ liệu gắn nhãn thủ công. |
| A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] | SfM cung cấp thông tin sơ bộ về vị trí camera và một mô hình điểm 3D thô.  MVS tinh chỉnh và tăng cường chi tiết dựa trên các dữ liệu này, dẫn đến mô hình hoàn chỉnh hơn. |
| Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] | Kết hợp dữ liệu có nhãn và không nhãn để huấn luyện mô hình độ sâu từ ảnh đơn (Monocular Depth Estimation - MDE). |

**2.6.2. Data**

|  |  |
| --- | --- |
| **Công trình** | **Data** |
| Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] | Sử dụng bộ dữ liệu ảnh đa thời gian tại nơi nghiên cứu thực địa  Sử dụng máy ảnh Nikon D7000 với ống kính cố định 20mm, khẩu độ f/2.8.  Máy ảnh gắn vào khung balo, kết nối với iPad qua bộ định tuyến không dây. |
| Photogrammetric point cloud generation and surface interpolation for change detection [2016] | 18 ảnh được chụp trên không dưới dạng **.tif**  Dữ liệu định hướng với định dạng **.ori** (chứa thông tin về ID ảnh, tọa độ máy bay, độ cao bay, và các tham số biến đổi cho mỗi ảnh) |
| Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] | Dữ liệu huấn luyện: Các đường viền 2D được tạo ra tự động bằng Grasshopper, với các biến thể về hình dạng, góc, độ cong, nhiễu, v.v.  Dữ liệu kiểm thử: Dữ liệu trắc địa ảnh thực tế (các đường viền của các tòa nhà) được sử dụng để kiểm tra hiệu quả của phương pháp. |
| A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] | 288 ảnh chụp cây táo từ nhiều góc nhìn, độ cao |
| Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] | Dữ liệu có nhãn:  Gồm 1.5 triệu ảnh từ 6 tập công khai như:  BlendedMVS: 115K ảnh, thu thập từ stereo.  MegaDepth: 128K ảnh, thu thập từ SfM.  DIML: 927K ảnh, bao gồm cảnh trong nhà và ngoài trời.  Nhãn độ sâu được thu thập từ các kỹ thuật stereo hoặc SfM.  Dữ liệu không nhãn:  62 triệu ảnh từ 8 tập dữ liệu công khai, như:  BDD100K: 8.2 triệu ảnh (cảnh giao thông).  Open Images V7: 7.8 triệu ảnh đa dạng.  SA-1B: 11.1 triệu ảnh từ nhiều ngữ cảnh khác nhau.  Các nhãn giả được tạo tự động bằng mô hình giáo viên. |

**2.6.3. Kết quả**

|  |  |
| --- | --- |
| **Công trình** | **Kết quả** |
| Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] | Độ chính xác cảu phương pháp quan trắc là 1.29mm  Sai số đứng (relative vertical accuracy) là 0.916 mm, rất nhỏ so với các kỹ thuật đo lường khác.  Sai số đứng tuyệt đối là ±8.26 cm, cho thấy khả năng sử dụng trong các nghiên cứu thực địa. |
| Photogrammetric point cloud generation and surface interpolation for change detection [2016] | Cả hai phần mềm ERDAS Imagine và Spacemetric Keystone đều thành công trong việc tạo ra đám mây điểm 3D với mật độ cao, bám sát cấu trúc bề mặt tổng thể.  Và vì áp dụng kỹ thuật lọc nhiễu và loại bỏ điểm ngoại lai có thể cải thiện đáng kể độ chính xác của DSM (Digital Surface Model), đặc biệt là ở khu vực rừng. |
| Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] | Phương pháp có khả năng phát hiện góc chính xác trên các đường viền khác nhau, kể cả các đường viền phức tạp và có nhiễu.  Mô hình được đào tạo bằng hệ thống phòng tập cho kết quả tốt trên dữ liệu thực tế.  Số lượng mặt của mô hình giảm đáng kể so với các phương pháp khác, giúp đơn giản hóa mô hình hiệu quả.  Phương pháp có thể tái tạo các bề mặt 3D một cách chính xác và dễ dàng tùy chỉnh. |
| A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] | Đường kính: Sai số tuyệt đối trung bình là 1.68 mm  Chiều dài: Sai số tuyệt đối trung bình là 1.35 mm  Sau khi điều chỉnh mô hình 3D để phản ánh sự co rút do mất nước, kết quả cho thấy sai số giảm xuống còn 0.29 mm cho đường kính và 1.32 mm cho chiều dài. |
| Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] | 1. So sánh với MiDaS (DPT-BEiTL-512):  - KITTI:  + AbsRel: giảm từ 0.127 → 0.076.  + 𝛿1: tăng từ 0.850 → 0.947.  - NYUv2:  + AbsRel: giảm từ 0.048 → 0.043.  + 𝛿1: tăng từ 0.980 → 0.981.  2. Fine-tuning trên độ sâu đo lường:  - NYUv2:  + AbsRel: giảm từ 0.069 → 0.056.  + 𝛿1: tăng từ 0.964 → 0.984.  - KITTI:  + AbsRel: giảm từ 0.054 → 0.046.  + 𝛿1: tăng từ 0.971 → 0.982.  3. Phân đoạn ngữ nghĩa:  - Cityscapes: mIoU tăng từ 82.1 (MiDaS) → 86.2 (Depth Anything).  - ADE20K: mIoU tăng từ 52.4 (MiDaS) → 59.4 (Depth Anything). |

**2.6.4. Ưu điểm**

|  |  |
| --- | --- |
| **Công trình** | **Ưu điểm** |
| Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] | Độ chính xác cao: Sử dụng SIFT cùng nội suy dữ liệu trên lưới 5mm giúp gia tăng độ chính xác, cải thiện độ chính xác của bộ dữ liệu DEM  Chi phí và hiệu quả: Quan trắc cự li gần giúp giảm chi phí hơn so với các phương pháp khác như LiDAr |
| Photogrammetric point cloud generation and surface interpolation for change detection [2016] | - Chi phí thấp: So với LiDAR(Light Detection and Ranging), việc thu thập ảnh chụp từ trên không có chi phí thấp hơn đáng kể.  - Độ phân giải cao: Ảnh chụp từ trên không thường có độ phân giải cao hơn dữ liệu LiDAR, cho phép tạo ra DSM chi tiết hơn.  - Dễ dàng thu thập dữ liệu: Việc thu thập ảnh chụp từ trên không dễ dàng và nhanh chóng hơn so với LiDAR, đặc biệt ở các khu vực rộng lớn. |
| Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] | Hiệu quả: Tái tạo mô hình một cách nhanh chóng và đơn giản.  Đơn giản: Sử dụng dữ liệu đơn giản (góc ngoài) thay vì hình ảnh phức tạp.  Tự động: Dữ liệu huấn luyện được tạo tự động, không cần gắn nhãn thủ công.  Linh hoạt: Có thể tùy chỉnh các tham số, hỗ trợ thiết kế tham số.  Khả năng xử lý nhiễu: Có thể xử lý các đường viền có nhiễu.  Tính trực quan: Mô hình được tái tạo gọn gàng, dễ sử dụng, dễ xuất sang các định dạng khác.  Ứng dụng đa dạng: Có tiềm năng sử dụng trong nhiều lĩnh vực, như kiến trúc, phim ảnh, trò chơi. |
| A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] | - Độ chính xác cao: Phương pháp cho phép tạo ra mô hình 3D chính xác, có thể sử dụng làm cơ sở để kiểm tra các hệ thống cảm biến khác.  - Không phá hủy: Kỹ thuật này không gây hại cho cây, cho phép theo dõi sự thay đổi theo thời gian.  - Phát hiện sự thay đổi: Có thể phát hiện và đo lường sự co rút của cây do mất nước. |
| Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] | - Khả năng tổng quát hóa mạnh mẽ trên các tập dữ liệu chưa từng thấy trước.  - Sử dụng dữ liệu không nhãn quy mô lớn, tiết kiệm chi phí thu thập.  - Encoder đa nhiệm: Hoạt động tốt trên cả độ sâu và phân đoạn ngữ nghĩa.  - Kết hợp nhiễu mạnh và hỗ trợ ngữ nghĩa, tăng khả năng học biểu diễn bền vững. |

**2.6.5. Khuyết điểm**

|  |  |
| --- | --- |
| **Công trình** | **Khuyết điểm** |
| Validation and application of close-range photogrammetry to quantify ephemeral gully erosion [2015] | Do hạn chế của việc sử dụng Trắc địa ảnh tầm gần mà quy mô thực hiện bị giới hạn  Tồn tại sai số của các thiết bị hỗ trợ, cũng như khả năng bị nhiễu bởi mỗi trường  Yêu cầu kĩ thuật cao, đòi hỏi người thực hiện phải đáp ứng đủ các quy trình thực hiện phức tạp |
| Photogrammetric point cloud generation and surface interpolation for change detection [2016] | - Độ chính xác thấp hơn LiDAR: Độ chính xác của DSM tạo ra từ so khớp ảnh thường thấp hơn so với LiDAR, đặc biệt ở khu vực rừng.  - Ảnh hưởng của điều kiện thời tiết: Mây che phủ, bóng râm, và điều kiện ánh sáng có thể ảnh hưởng đến chất lượng của so khớp ảnh.  - Khó khăn trong xử lý nhiễu: Đám mây điểm từ so khớp ảnh thường chứa nhiều nhiễu, đặc biệt ở khu vực rừng, đòi hỏi kỹ thuật xử lý phức tạp.  - Yêu cầu dữ liệu định hướng chính xác: Dữ liệu định hướng chính xác cho mỗi ảnh là rất quan trọng để đảm bảo độ chính xác của kết quả. |
| Reconstructing Photogrammetric 3D Model by Using Deep Learning [2021] | Vẫn có thể có sai sót nhỏ: Vẫn còn một số lỗi nhỏ có thể xảy ra xung quanh các góc, tạo ra các phân đoạn nhỏ.  Độ phân giải góc còn hạn chế: Độ phân giải phát hiện góc hiện tại là 100 nút, có thể không đủ trong một số trường hợp đặc biệt.  Vấn đề texture chưa được giải quyết: Phương pháp này chưa giải quyết được vấn đề kết cấu (texture) của các bề mặt.  Chưa có kiểm định toàn diện: Chưa có nhiều độ đo và so sánh chi tiết với các phương pháp khác. |
| A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees [2021] | - Thời gian thực hiện: Quy trình tạo ra mô hình 3D có thể tốn thời gian và yêu cầu nhiều công sức.  - Yêu cầu điều kiện ánh sáng tốt: Kết quả có thể bị ảnh hưởng bởi điều kiện ánh sáng không phù hợp.  - Độ chính xác phụ thuộc vào kỹ thuật: Nếu không thực hiện đúng quy trình, độ chính xác của mô hình có thể bị giảm. |
| Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data [2024] | - Phụ thuộc vào mô hình tiền huấn luyện: Khả năng của mô hình bị giới hạn bởi hiệu suất của DINOv2 (mô hình tiền huấn luyện).  - Độ phân giải huấn luyện thấp: Hiện tại chỉ huấn luyện ở độ phân giải 512×512, chưa tận dụng được hết tiềm năng.  - Chi phí tính toán lớn: Xử lý dữ liệu không nhãn quy mô lớn đòi hỏi tài nguyên tính toán cao.  - Nhiễu trong nhãn giả: Nhãn giả từ mô hình giáo viên có thể mang theo lỗi, ảnh hưởng đến hiệu suất học tập của mô hình sinh viên. |

# **Chương 3. Giải pháp [10]**

Do cấu hình thiết bị của nhóm chúng em bị hạn chế và việc tìm hiểu các thư viện Python chưa đủ sâu rộng, nên sau nhiều thử nghiệm tự xây dựng chương trình thất bại và không thu được kết quả mong muốn, nhóm chúng em đã quyết định sử dụng lại một số công trình mà những người đi trước đã nghiên cứu và cài đặt: Meshroom.

Meshroom là một ứng dụng được xây dựng dựa trên kiến trúc quang trắc (photogrammetry pipeline) mà nhóm AliceVision, một tổ chức phi lợi nhuận về phát triển các sản phẩm thị giác máy tính ba chiều tại châu Âu, đã nghiên cứu và phát triển từ năm 2015. Kiến trúc quang trắc của ứng dụng hiện tại đã được hoàn thành vào năm 2021 và cải tiến từ đó đến nay. Dưới đây là chi tiết về kiến trúc:

## **3.1.** **Trích xuất các đặc trưng bất biến**

Tác giả áp dụng phương pháp SIFT để tìm kiếm và triết suất các đặc trưng cố định trên bộ ảnh. Cụ thể, phương pháp này sẽ được thực hiện theo các bước sau:

* Tạo "pyramid" (kim tự tháp) của ảnh:
  + Để nhận diện các điểm đặc trưng ở nhiều kích thước khác nhau, ảnh gốc được làm mờ và thu nhỏ dần để tạo một chuỗi các phiên bản ảnh (gọi là "pyramid").
  + Mỗi tầng của pyramid chứa ảnh với kích thước nhỏ hơn tầng phía trên.
* Tính toán DoG (Difference of Gaussians)
  + Phương pháp này giúp tìm ra các điểm mang nhiều thông tin nhất trong ảnh, gọi là các điểm cực đại (maxima) hoặc cực tiểu (minima).
* Xác định Keypoints
  + Tại mỗi điểm cực đại/cực tiểu tìm được, SIFT trích xuất một vùng vuông nhỏ (patch) xung quanh điểm đó.
  + Trong vùng này, SIFT tính toán hướng gradient thống trị (gradient chính) để ghi nhận độ xoay của điểm đặc trưng.
* Tạo mô tả cho từng keypoint
  + Mỗi keypoint sẽ được mô tả bằng một vector chứa thông tin về hướng và cường độ gradient tại vùng lân cận của nó.
  + Vector mô tả này thường có 128 giá trị, tương ứng với thống kê các hướng gradient trong khu vực xung quanh keypoint.
  + Đặc điểm này giúp SIFT nhận diện keypoints ngay cả khi hình ảnh bị xoay hoặc thay đổi.
* Lọc kết quả
  + Vì số lượng keypoints trích xuất được phụ thuộc vào độ phức tạp của bề mặt ảnh, cần có bước lọc để giới hạn số lượng keypoints (ví dụ, từ vài chục đến vài nghìn keypoints mỗi ảnh).
  + Lọc này đảm bảo các đặc trưng được phân bố đều trên toàn bộ ảnh.

## **3.2.** **So khớp ảnh**

Tác giả sử dụng Cây từ vựng (Vocabulary Tree) để đơn giản hóa kỹ thuật:

* Tất cả các đặc trưng đã được trích xuất từ ảnh sẽ được so sánh với các đặc trưng tại mỗi nút của cây từ vựng.
* Mỗi đặc trưng sẽ rơi vào một lá cụ thể trong cây, và lá đó được gắn với một chỉ mục đơn giản (index).
* Biểu diễn đặc trưng của hình ảnh sau đó được thể hiện dưới dạng tập hợp các chỉ mục của các lá mà hình ảnh đó sử dụng.

Thay vì phải kiểm tra từng đặc trưng chi tiết, ta có thể dễ dàng so sánh và xác định xem hai hình ảnh khác nhau có chia sẻ cùng nội dung hay không thông qua việc so sánh các chỉ mục đặc trưng này.

## **3.3.** **So khớp đặc trưng**

Mục tiêu của bước này là tìm ra tất cả các cặp đặc trưng tương ứng giữa các cặp hình ảnh ứng viên. Các bước thực hiện bao gồm:

* Tạo danh sách ứng viên: Đối với mỗi đặc trưng được trích suất trong hình ảnh A, tìm các đặc trưng tương tự từ hình ảnh B bằng cách so sánh các descriptor.
* Lọc các ứng viên không hợp lệ:
  + Tìm 2 đặc trưng gần nhất từ hình ảnh B cho mỗi đặc trưng trong hình ảnh A (Sử dụng thuật toán Approximate Nearest Neighbor (ANN) hoặc Cascading Hashing để tăng tốc độ tìm , thay vì sử dụng phương pháp kiểm tra toàn diện.)
  + Sử dụng ngưỡng tương đối để loại bỏ các đặc trưng trùng lặp hoặc không phù hợp.
* Lọc bằng hình học không gian:
  + Áp dụng phương pháp RANSAC (RANdom SAmple Consensus) để kiểm tra sự phù hợp của các cặp đặc trưng dựa trên ma trận cơ bản (fundamental hoặc essential matrix).
  + Giữ lại các đặc trưng phù hợp nhất sau khi kiểm tra bằng RANSAC.

## **3.4. Tái tạo cấu trúc (Structure from Motion)**

* Hợp nhất đặc trưng:
  + Kết hợp các đặc trưng được khớp giữa các cặp hình ảnh để tạo thành các đường track, đại diện cho các điểm trong không gian.
  + Loại bỏ các track không phù hợp (outliers).
* Chọn cặp hình ảnh khởi đầu:
  + Chọn cặp hình ảnh đầu tiên có nhiều đặc trưng khớp và cung cấp thông tin hình học đáng tin cậy nhất.
  + Góc giữa các máy ảnh trong cặp này cần đủ lớn để cung cấp dữ liệu hình học chất lượng.
* Tính toán fundamental matrix:
  + Xác định fundamental matrix giữa hai hình ảnh đầu tiên.
  + Sử dụng ma trận này để xác định tư thế (pose) của hai camera đầu tiên và suy ra các điểm 3 chiều từ các đặc trưng 2 chiều bằng phương pháp tam giác hóa.
* Chọn các hình ảnh tiếp theo:
  + Sử dụng thuật toán Next Best View Selection để chọn hình ảnh tiếp theo có đủ đặc trưng liên kết với các điểm 3D đã tái tạo.
  + Dự đoán tư thế của các camera mới bằng thuật toán PnP (Perspective-n-Point) trong khung RANSAC và tinh chỉnh bằng tối ưu hóa phi tuyến.
* Tinh chỉnh toàn bộ:
  + Tinh chỉnh các tham số bên ngoài (extrinsic), bên trong (intrinsic) của camera và vị trí của các điểm 3 chiều.
  + Loại bỏ các quan sát có lỗi tái chiếu (reprojection error) cao hoặc góc quan sát không đủ lớn.

Lặp lại quy trình trên cho đến khi tất cả các hình ảnh được sử dụng hoặc không thể thêm các điểm 3 chiều mới.

## **3.5. Ước lượng bản đồ độ sâu (Depth Map Estimation)**

* Lựa chọn camera lân cận:
  + Với mỗi ảnh, chọn N camera gần nhất đã được xác định từ quá trình SfM.
  + Xác định các mặt phẳng song song với trục quang học của camera, tạo một không gian 3 chiều (chiều rộng, chiều dài, chiều cao) với nhiều giá trị độ sâu ứng viên cho mỗi điểm ảnh.
* Tính toán độ tương đồng:
  + Đo lường độ tương đồng (similarity) giữa các giá trị ứng viên bằng Zero Mean Normalized Cross-Correlation (ZNCC) trên từng vùng ảnh nhỏ (patch).
  + Tạo một khối lượng dữ liệu (volume) chứa các giá trị độ tương đồng giữa các ảnh.
* Lọc và giảm nhiễu:
  + Lọc theo các trục X và Y để giảm giá trị nhiễu cục bộ.
  + Chọn giá trị độ sâu tối thiểu cục bộ (local minima) từ dữ liệu lọc để xây dựng bản đồ độ sâu.
* Tinh chỉnh độ sâu:
  + Áp dụng bước tinh chỉnh (refine) để cải thiện độ chính xác của giá trị độ sâu ở cấp độ sub-pixel.
  + Loại bỏ các hiện tượng "banding" gây ra bởi việc chọn giá trị độ sâu ban đầu.
* Đảm bảo tính nhất quán giữa các camera:
  + Tính toán bản đồ độ sâu cho từng ảnh độc lập (có thể thực hiện song song).
  + Áp dụng bước lọc cuối cùng để đảm bảo tính nhất quán giữa các bản đồ độ sâu từ nhiều camera, cân bằng giữa giá trị độ tương đồng và số lượng bề mặt được hỗ trợ.

## **3.6. Tạo khung lưới cho mô hình (Meshing)**

* Hợp nhất bản đồ độ sâu: Tất cả các bản đồ độ sâu (depth maps) được hợp nhất vào một cấu trúc octree toàn cục, nơi các giá trị độ sâu tương thích được ghép vào các ô (cells) của octree.
* Tạo khối tứ diện 3 chiều:
  + Sử dụng thuật toán Delaunay tetrahedralization để chia không gian thành các khối tứ diện.
  + Thực hiện một quy trình complex voting để tính toán trọng số cho các ô và các mặt kết nối giữa chúng.
* Tối ưu hóa bề mặt:
  + Áp dụng thuật toán Graph Cut Max-Flow để cắt tối ưu khối lượng, trích xuất bề mặt lưới (mesh surface).
  + Lọc bỏ các ô không phù hợp trên bề mặt.
* Làm mịn lưới: Áp dụng lọc Laplacian để loại bỏ các hiện tượng nhiễu cục bộ và làm mịn bề mặt lưới.
* Đơn giản hóa lưới: Lưới có thể được đơn giản hóa thêm bằng cách giảm số lượng đỉnh không cần thiết để tối ưu hóa cấu trúc.

## **3.7. Khớp các chi tiết vào lưới (Texturing)**

* Tạo bản đồ UV tự động (UV Mapping):
  + Nếu lưới chưa có bản đồ UV, hệ thống sẽ tự động tính toán bản đồ UV để tối ưu hóa không gian texture.
  + Sử dụng phương pháp UV Mapping cơ bản [8]
* Thu thập dữ liệu texture từ các camera:
  + Với mỗi tam giác trên lưới, sử dụng thông tin hiển thị (visibility) từ các đỉnh (vertex) để lấy các texture phù hợp từ các camera.
  + Loại bỏ các camera không có góc nhìn tốt đối với bề mặt tam giác.
* Tổng hợp giá trị pixel:
  + Thay vì chỉ trung bình đơn giản (naive averaging), áp dụng phương pháp tổng hợp đa dải tần (multi-band blending). [9]
  + Các chi tiết tần số thấp được tổng hợp từ nhiều góc nhìn hơn so với chi tiết tần số cao để tăng độ chính xác và tự nhiên.
* Kết quả cuối cùng: Lớp texture hoàn thiện được áp dụng lên lưới 3 chiều, tạo ra bề mặt với màu sắc và chi tiết trực quan.

# **Chương 4. Cài đặt và thử nghiệm chương trình**

Do chương trình được xây dựng dựa trên các phương pháp truyền thống, nên không có giai đoạn huấn luyện.

Chương trình được chạy thử nghiệm với cấu hình máy như sau:

+ CPU: Intel Core i5-12500H

+ GPU: NVIDIA GeForce RTX 3050

+ RAM: 16 GB

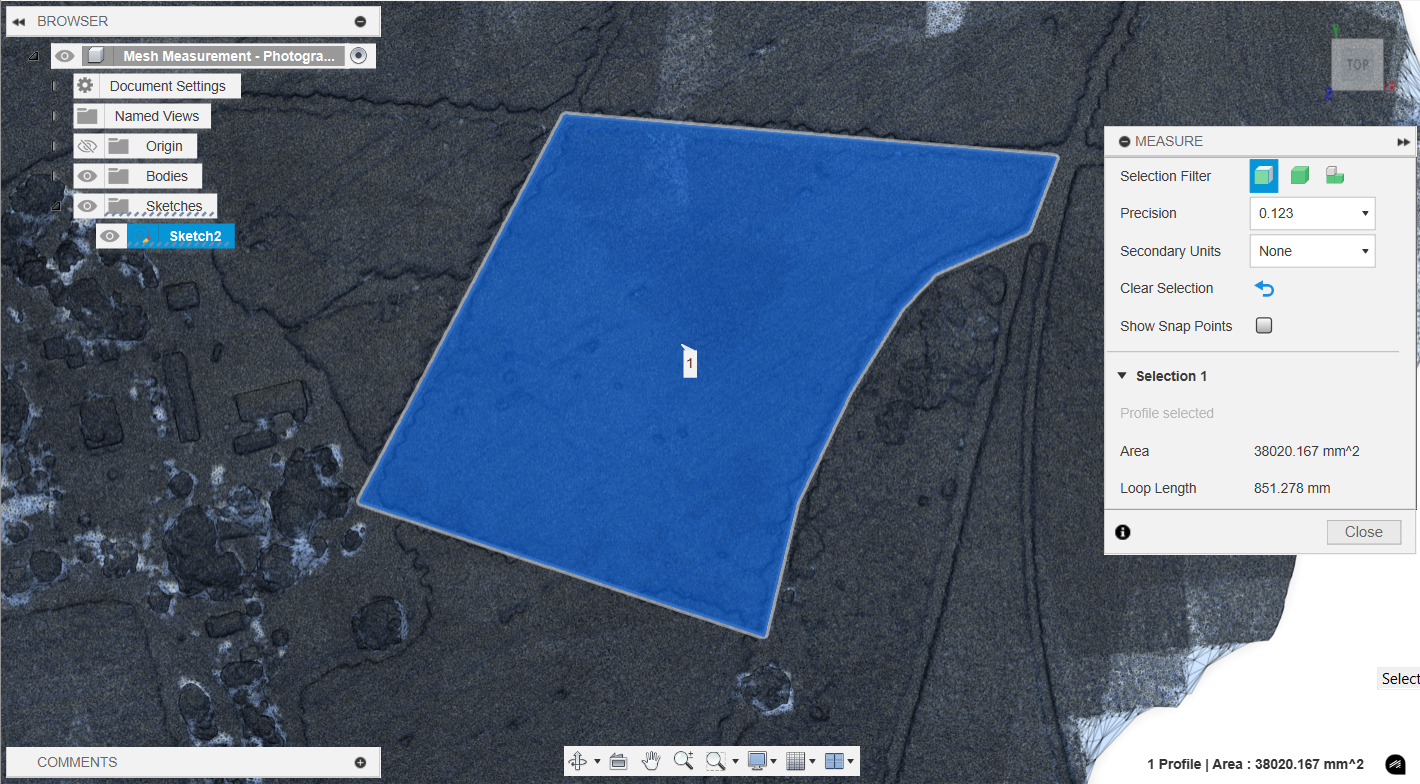
Tập dữ liệu mẫu chúng em sử dụng để chạy thử được tải xuống từ đường dẫn này: <https://s3.eu-central-1.amazonaws.com/ageagle.com/Dataset+-+Corn+field+eBee+Ag/RGB%2BImages.zip>

Nhóm chúng em đã lấy 67 ảnh có tên: ‘IX-12-85231\_0010\_0001.JPG’, ‘IX-12-85231\_0010\_0002.JPG’, …, ‘IX-12-85231\_0010\_0067.JPG’ để thử nghiệm với chương trình. Dưới đây là ảnh kết quả của thử nghiệm đó:

A map of a field

Description automatically generated

Sau khi hoàn thành công đoạn tái tạo, mô hình 3 chiều này có thể được xuất thành 1 tập tin .obj và được sử dụng cho việc đo đạc hoặc mô phỏng lại. Trước khi tiến hành đo đạc, nhóm đã di chuyển và xoay bề mặt để bề mặt nằm trên một mặt phẳng tọa độ. Dưới đây là ảnh đo đạc mà nhóm chúng em đã thực hiện trên AutoCAD:



Hiện tại nhóm chúng em chưa thể đo diện tích của khu vực này theo đúng tỉ lệ, do chưa khai thác triệt để metadata của tập dữ liệu mẫu.

# **Chương 5: Hạn chế và hướng phát triển**

## **5.1. Hạn chế**

Do đây là chương trình được xây dựng theo mã nguồn cho trước và sử dụng nhiều phương pháp truyền thống nên kết quả của chương trình chưa thật sự đạt được tới mức SOTA. Mặt khác, do hạn chế về mặt thiết bị, nhóm chúng em chỉ có thể chạy trên một số lượng ảnh nhất định.

## **5.2. Phương hướng phát triển**

Dưới đây là phương hướng phát triển mà nhóm chúng em muốn hướng tới:

* Viết lại chương trình theo thư viện và ngôn ngữ lập trình Python: Với ràng buộc từ việc xây dựng ứng dụng trên mã nguồn có sẵn, vấn đề cải tiến chương trình đã gặp khó khăn. Do đó, nhóm chúng em có định hướng viết lại kiến trúc của AliceVision theo ngôn ngữ Python, tích hợp với một số thư viện mã nguồn mở dành cho Python.
* Thay thế một số phương pháp truyền thống ở một số bước bằng các phương pháp tích hợp học sâu: Sau khi viết lại chương trình, chúng em định hướng sẽ thay thế một số phương pháp truyền thống bằng các phương pháp tiên tiến hơn, từ đó thực hiện so sánh và phát triển các bước trong kiến trúc để kết quả đạt được tốt hơn.
* Tích hợp thêm các tính năng liên quan đến đo đạc, phát hiện vật thể, v.v…: Điều này giúp đa dạng hóa tính năng của chương trình, đồng thời là tiền đề để phát triển các ứng dụng thực tế trong nông nghiệp hoặc kiến trúc.

# **Chương 6: Tham khảo**

[1] THE APPLICATION OF UAV AND PHOTOGRAMMETRY FOR SUPPORTING PRECISION AGRICULTURE AND MONITORING ENVIRONMENTAL PROBLEMS (2016)

[2] Elements of Photogrammetry with Applications in GIS Paul R. Wolf, Ph.D. Bon A. Dewitt, Ph.D. Benjamin E. Wilkinson, Ph.D. (2014)

[3] A photogrammetry-based methodology to obtain accurate digital ground-truth of leafless fruit trees Bernat Lavaquiol \* , Ricardo Sanz , Jordi Llorens, Jaume Arno ́ , Alexandre Escola` 2021

[4] Photogrammetric point cloud generation and surface interpolation for change detection (2016)

[5] Reconstructing Photogrammetric 3D Model by Using Deep Learning, June-Hao Hou, Chili Cheng (2021)

[6] Validation and application of close-range photogrammetry to quantify ephemeral gully erosion (2015)

[7] Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data, Lihe Yang Bingyi Kang, Zilong Huang, Xiaogang Xu(2024)

[8] Least Squares Conformal Maps for Automatic Texture Atlas Generation, Bruno Lévy, Sylvain Petitjean, Nicolas Ray, Jérome Maillot, Siggraph 2002

[9] A Multiresolution Spline with Application to Image Mosaics, P. J. Burt and E. H. Adelson, ACM Trans. Graph. 1983

[10] [AliceVision Meshroom: An open-source 3D reconstruction pipeline](https://alicevision.org/), Carsten Griwodz, Simone Gasparini, Lilian Calvet, Pierre Gurdjos, Fabien Castan, Benoit Maujean, Gregoire De Lillo, Yann Lanthony, 2021