

# CƠ SỞ KHOA HỌC PHƯƠNG PHÁP ĐO ẢNH CHỤP TỪ THIẾT BỊ BAY KHÔNG NGƯỜI LÁI UAV

NGUYỄN ĐẠI ĐỒNG<sup>(1)</sup>, BÙI NGỌC QUÝ<sup>(2)</sup>

<sup>(1)</sup>Cục Đo đạc, Bản đồ và Thông tin địa lý Việt Nam

<sup>(2)</sup>Trường Đại học Mở - Địa chất

## **Tóm tắt:**

Bài báo trình về cơ sở khoa học của phương pháp đo ảnh chụp từ thiết bị bay không người lái UAV dựa trên giải thuật biến đổi đối tượng phi tỷ lệ (Scale-Invariant Feature Transform - SIFT) và phương pháp tái tạo hình ảnh dựa vào chuyển động (Structure from Motion - SfM) trong công tác thành lập bản đồ địa hình.

## **1. Đặt vấn đề**

Công nghệ đo vẽ ảnh (Photogrammetry) đã được sử dụng quen thuộc từ trước đến nay. Tuy nhiên với sự phát triển nhanh chóng của công nghệ, đặc biệt là công nghệ thông tin, máy tính; các phương pháp đo vẽ ảnh lập thể truyền thống 3D, sau khi tăng dày và khôi phục mô hình lập thể đã dần được thay thế bằng các phương pháp đo vẽ tự động trên máy tính dựa trên nền tảng của giải thuật biến đổi đối tượng phi tỷ lệ (Scale Invariant Feature Transform) và phương pháp tái tạo hình ảnh dựa vào chuyển động (Structure from Motion). Giải thuật biến đổi đối tượng phi tỷ lệ SIFT sử dụng trong lĩnh vực thị giác máy tính (Computer Vision) dùng để nhận dạng và miêu tả những điểm đặc trưng (local features) trong ảnh, giải thuật này phát hiện và tạo một tập hợp số lượng rất lớn các điểm đặc trưng. Giải thuật SfM cũng tiến hành nhận diện các điểm đặc trưng (Key Points) và sau đó được sử dụng để tạo ảnh lập thể 3D từ các ảnh 2D. Các giải thuật SIFT và SfM được ứng dụng nhiều trong các phần mềm xử lý dữ liệu ảnh chụp UAV phổ biến hiện nay.

## **2. Cơ sở khoa học phương pháp đo ảnh từ thiết bị bay không người lái UAV**

### **2.1. Tổng quan về phương pháp đo ảnh UAV**

Phương pháp bình sai khối ảnh (Adjustment

Bundle Block BBA) được sử dụng trong đo ảnh hàng không truyền thống, để xác định các yếu tố định hướng ngoài của mỗi tấm ảnh. Thông thường trong chụp ảnh hàng không các thông số định hướng trong của ảnh (bao gồm độ dài tiêu cự, tọa độ điểm chính ảnh, các thông số méo ảnh) được biết trước. Tuy nhiên, đối với máy ảnh gắn trên các thiết bị UAV thì các thông số này thường không được biết trước hoặc không ổn định bởi các máy chụp ảnh này, thường là máy chụp ảnh phổ thông.

Ngày nay, trên các thiết bị bay chụp thường được tích hợp với các hệ thống định vị vệ tinh GNSS và hệ thống dẫn đường quán tính INS cho phép xác định vị trí chụp ảnh và các góc xoay với độ chính xác cao, điều này giúp tham chiếu ảnh (Georeferencing) trực tiếp mà không cần tới điểm khống chế ảnh ngoại nghiệp (GCP). Thông thường các điểm GCP được sử dụng để đảm bảo độ tin cậy khi chuyển tọa độ GNSS về hệ tọa độ địa phương.

Trong bình sai khối ảnh, yêu cầu bắt buộc phải có các điểm song trùng (điểm ảnh cùng tên) của các tấm ảnh phủ lên nhau. Nhờ thuật toán SIFT, các điểm này sẽ được phát hiện và tạo một tập hợp số lượng rất lớn giúp nâng cao độ chính xác của kết quả thu được trong đo vẽ ảnh UAV.

Do các tham số định hướng trong và định hướng ngoài thường không biết trước khi đo vẽ

ảnh UAV do đó gây ra vấn đề trong bình sai khối ảnh truyền thống. Để giải quyết vấn đề này, dữ liệu GPS và IMU (thiết bị đo quán tính) đã được sử dụng để ước lượng các tham số định hướng ngoài gần đúng, sau đó đưa vào các phần mềm xử lý ảnh số để khớp ảnh trong nghiên cứu của Berni và nnk (2009) [2]. Sau đó các điểm GCP được sử dụng để tính toán tam giác ảnh không gian. Cuối cùng DTM sẵn có được sử dụng để tạo ảnh trực giao.

Về lý thuyết, ảnh chụp từ UAV có nguyên lý ghép khối cũng tương tự như nguyên lý ghép khối ảnh chụp lập thể từ máy bay có người lái truyền thống. Tuy nhiên, số lượng ảnh UAV cần thiết để bao phủ khu vực cần chụp ảnh thường rất lớn từ vài chục đến vài trăm, thậm trí là hàng nghìn tấm ảnh do ảnh UAV thường có kích thước nhỏ vì được chụp ở độ cao thấp. Do đó sẽ không hiệu quả hoặc không thể sử dụng được phương pháp truyền thống. Đặc biệt là khâu chọn các điểm song trùng hay tăng dày không chế ảnh nội nghiệp. Nhiều nhà khoa học trên thế giới để giải quyết vấn đề này, đã nghiên cứu phương pháp xử lý ảnh chiết tách các điểm song trùng, xây dựng đám mây điểm dày đặc và ảnh trực giao bằng việc áp dụng các thuật toán khớp ảnh tự động. Một trong những phương pháp tiêu biểu và được sử dụng phổ biến hiện nay được gọi tắt là SfM [10]. Phương pháp này sử dụng các điểm đặc trưng là kết quả của giải thuật biến đổi đối tượng phi tỉ lệ (Scale-Invariant Feature Transform - SIFT).

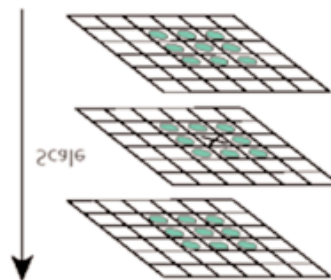
## 2.2. Giải thuật biến đổi đối tượng phi tỉ lệ (Scale-Invariant Feature Transform - SIFT)

Bản chất của giải thuật biến đổi đối tượng phi tỉ lệ (Scale-Invariant Feature Transform - SIFT) chính là phương pháp trích rút các đặc trưng cục bộ bất biến SIFT của ảnh. Các đặc trưng này đối với việc thay đổi tỉ lệ ảnh, xoay ảnh, đôi khi là thay đổi điểm nhìn và thêm nhiều ảnh hay thay đổi cường độ chiếu sáng của ảnh là bất biến. Phương pháp được lựa chọn có tên là Scale-Invariant Feature Transform (SIFT) và đặc trưng trích rút được gọi là đặc trưng SIFT (SIFT Feature). Các đặc trưng SIFT này được trích rút

ra từ các điểm đặc trưng cục bộ (Local Interest Point) [7][4]. Các vị trí (điểm ảnh) có đặc điểm “đặc trưng” trên ảnh được gọi là điểm đặc trưng (Interest Point (Keypoint)). Sự “đặc trưng” ở đây có nghĩa là điểm đó có thể có các đặc trưng bất biến với việc quay ảnh, co giãn ảnh hay thay đổi cường độ chiếu sáng của ảnh. Phương pháp trích xuất các đặc trưng bất biến SIFT được tiếp cận theo phương pháp thác lọc, theo đó phương pháp được thực hiện lần lượt theo các bước sau:

- Phát hiện các điểm cực trị Scale-Space (Scale-Space extrema detection):

Bước đầu tiên này tiến hành tìm kiếm các điểm đặc trưng trên tất cả các tỉ lệ và vị trí của ảnh. Nó sử dụng hàm Different-of-Gaussian để xác định tất cả các điểm đặc trưng tiềm năng mà bất biến với quy mô và hướng của ảnh. Một không gian tỷ lệ được xây dựng bằng cách tiến hành tích chập ảnh lập đi lập lại sử dụng phép lọc Gauss (gọi tắt là ảnh Gauss) với việc thay đổi tỷ lệ và nhóm kết quả vào các nhóm (Lowe, 2004). Tiếp theo đó điểm song trùng được xác định là điểm cực trị cục bộ của ảnh D qua các tỷ lệ khác nhau. Mỗi điểm ảnh trong D được so sánh với 8 điểm ảnh láng giềng của nó ở cùng tỉ lệ đó và 9 láng giềng kề ở các tỉ lệ ngay trước và sau nó. Nếu điểm ảnh đó đạt giá trị cực tiểu hoặc cực đại thì sẽ được chọn làm các điểm đặc trưng dự kiến.

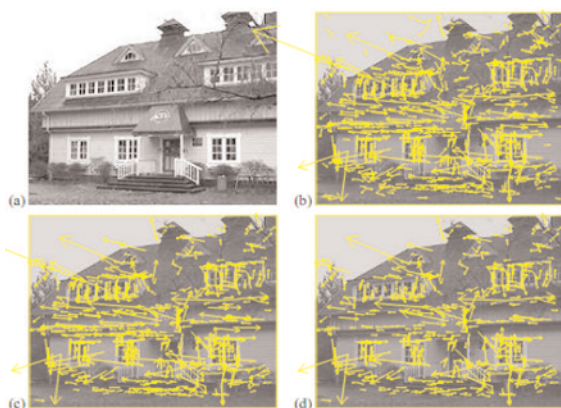


Hình 1: Mỗi điểm ảnh được so sánh với 26 láng giềng của nó

- Định vị các điểm đặc trưng (keypoint localization):

Mỗi ứng viên là điểm đặc trưng sau khi được

chọn sẽ được đánh giá xem có được giữ lại hay không. Một số điểm đặc trưng dọc theo các cạnh không giữ được tính ổn định khi ảnh bị nhiễu và xoay ảnh cũng bị loại bỏ.



Hình 2: Quá trình lựa chọn các điểm đặc trưng:

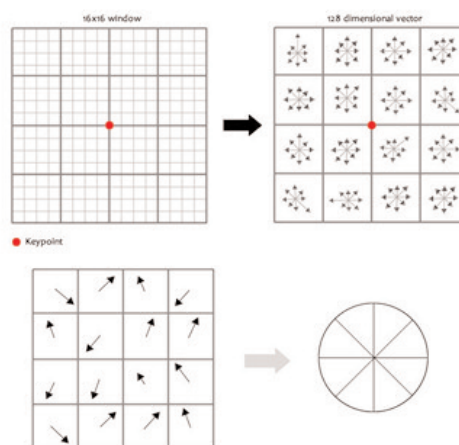
- Ảnh gốc,
- Các điểm đặc trưng được phát hiện,
- Ảnh sau khi loại bỏ các điểm đặc trưng có độ tương phản thấp,
- Ảnh sau loại bỏ các điểm đặc trưng dọc theo cạnh.

- Xác định hướng cho các điểm đặc trưng (Orientation assignment):

Điểm đặc trưng sau đó được định hướng dựa vào hướng dốc cục bộ (Lowe, 2004). Một tập hợp của hướng độ xám được tạo ra trên các điểm lân cận điểm đặc trưng.

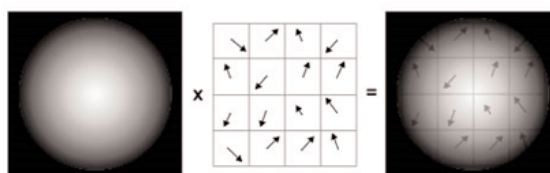
- Biểu diễn vector cho điểm đặc trưng (Keypoint descriptor):

Để xác định vector cho điểm đặc trưng, một cửa sổ  $16 \times 16$  sẽ được xác định xung quanh điểm đó. Cửa sổ  $16 \times 16$  này được chia thành mười sáu cửa sổ  $4 \times 4$ . Trong mỗi cửa sổ  $4 \times 4$ , độ lớn và hướng của vector được tính toán và đưa vào một biểu đồ 8 khối. Điểm đặc trưng sau khi được xác định hướng sẽ được biểu diễn dưới dạng các vector  $4 \times 4 \times 8 = 128$  chiều.



Hình 3: Biểu diễn các vector đặc trưng

Bất kỳ hướng dốc nào trong phạm vi  $0^\circ - 44^\circ$  sẽ thêm vào khối thứ nhất.  $45^\circ - 89^\circ$  thêm vào khối tiếp theo. Và như vậy số hướng được thêm vào các khối phụ thuộc vào độ lớn của hướng. Số hướng được thêm vào cũng phụ thuộc vào khoảng cách từ điểm đặc trưng. Vì vậy, càng ở xa điểm đặc trưng thì các giá trị được thêm vào biểu đồ sẽ càng nhỏ. Điều này được thực hiện bằng cách sử dụng “hàm trọng số gaussian”. Hàm này chỉ đơn giản là tạo độ dốc (giống như đường cong 2D). Chúng ta nhân nó với độ lớn của định hướng để có được một trọng số.



Hình 4: Mô tả hàm trọng số Gaussian cho hướng của điểm đặc trưng

Làm điều này cho tất cả 16 pixel, chúng ta sẽ “biên dịch” 16 hướng hoàn toàn ngẫu nhiên thành 8 khối được xác định trước và thực hiện cho tất cả 16 khu vực  $4 \times 4$  sau khi kết thúc ta sẽ được 128 hướng. Khi chúng ta có tất cả 128 hướng, tiến hành bình thường hóa chúng (giống như bình thường hóa một vector rồi chia cho góc của tổng bình phương). 128 hướng này tạo thành “vector đặc trưng”. Điểm đặc trưng này được xác định duy nhất bởi các vector đặc trưng này.

Có thể đã thấy rằng trong các hình ảnh trên, điểm đặc trưng nằm ở “giữa” và không nằm chính xác trong một pixel nào. 16x16 cửa sổ có các hướng và độ lớn của hình ảnh “ở giữa” pixel. Vì vậy, chúng ta cần nội suy hình ảnh để tạo dữ liệu định hướng và cường độ “ở giữa” pixel.

Trong công nghệ UAV, giải thuật biến đổi đối tượng phi tỷ lệ SIFT được sử dụng để thực hiện việc khớp các tấm ảnh UAV để tạo ra bình đồ ảnh. Đầu tiên là việc phân tích phủ chùm chuỗi hình ảnh và một số công tác tiền xử lý sau đó sử dụng thuật toán SIFT để trích xuất các điểm đặc trưng, tiếp đó lấy hướng chính của các điểm lân cận của điểm đặc trưng làm vector hướng. Tập hợp các điểm đặc trưng được trích xuất bởi thuật toán SIFT sau đó chúng ta có thể khớp chúng lại với nhau và tính toán mối quan hệ giữa các điểm ảnh lân cận (Neighbor points). Mối quan hệ giữa các điểm ảnh lân cận có thể được coi là hình chiếu phối cảnh, do đó biến đổi đồng nhất có thể được tính bằng:

$$\begin{cases} x' = \frac{m_1 x + m_2 y + m_3}{m_7 x + m_8 y + 1} \\ y' = \frac{m_4 x + m_5 y + m_6}{m_7 x + m_8 y + 1} \end{cases} \quad (1)$$

Trong đó: (x,y) là tọa độ điểm đặc trưng trên ảnh 1

(x', y') là tọa độ điểm đặc trưng tương ứng trên ảnh 2

$M=[m_1, m_2, m_3, \dots, m_8, 1]T$  là 8 tham số của phép biến đổi đồng nhất cũng có thể được biểu thị dưới dạng ma trận 3\*3 như sau:

$$\vec{M} = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & 1 \end{bmatrix} \quad (2)$$

Vì vậy chúng ta có cần phải sử dụng tối thiểu 4 điểm để tính toán ma trận biến đổi.

**Hiệu chỉnh độ cao máy bay:** Trong thực tế thiết bị UAV khi bay chụp không thể giữ trạng thái giống nhau hoàn toàn tại mọi thời điểm vì thế các vật thể tương tự trên các tấm ảnh lân cận có thể không được hiển thị với cùng hình dạng do đó kết quả khớp ảnh sẽ bị ảnh hưởng bởi loại

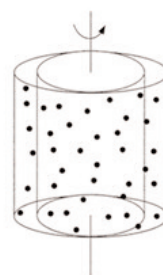
biến dạng này. Chính vì vậy, một số hiệu chỉnh nên được áp dụng trước khi tính toán mối quan hệ giữa từng cặp hình ảnh và thông tin độ cao bay có thể được cung cấp bởi Hệ thống định vị và định hướng trên UAV.

**Lựa chọn mặt phẳng tham chiếu:** Hình ảnh đầu tiên của dải bay thứ nhất có thể được chọn làm mặt phẳng tham chiếu trong quá trình khớp ảnh. Tuy nhiên, sai số méo hình có thể được tích lũy ngày càng lớn khi số lượng chuỗi hình ảnh ngày càng tăng. Do đó, nên chọn mặt phẳng tham chiếu thích hợp để giảm sai số méo hình tích lũy và biến dạng có thể được phân phối đều cho mỗi chuỗi hình ảnh.

Kết quả cuối cùng là bình đồ ảnh được khớp từ các tấm ảnh thành phần, vì vậy mặt phẳng tham chiếu có thể được chọn ngẫu nhiên.

### 2.3. Phương pháp tái tạo hình ảnh dựa vào chuyển động (Structure from Motion-SfM)

Nền tảng đầu tiên của phương pháp tái tạo hình ảnh dựa vào chuyển động SfM được đưa ra bởi S. Ullman. Theo đó, phương pháp tính toán vị trí 3D bằng các hình ảnh 2D được sử dụng để xác định SfM [3]. Hình chiếu phẳng và tọa độ 3D của các điểm ảnh được tính toán, đưa lên màn hình và được lưu lại trong bộ nhớ máy tính. Khi các hình trụ này xoay, các tọa độ và hình chiếu được tính toán lại để đưa lên mặt phẳng.

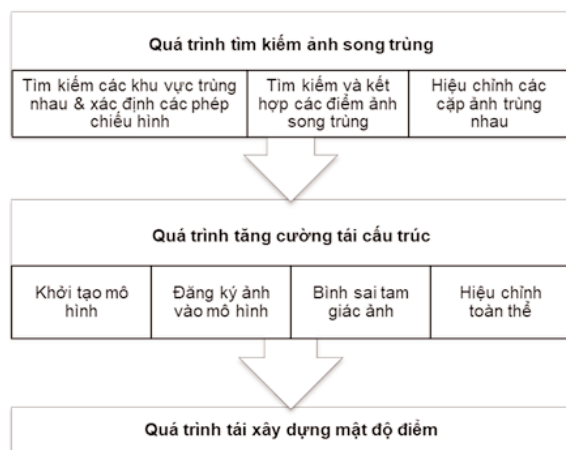


Hình 5: Cặp hình trụ miêu tả SfM.

Có thể nói, một tập điểm giống như ngẫu nhiên được tạo ra bởi mỗi hình ảnh trên màn hình sẽ xuất hiện. Nhưng, các đối tượng dịch chuyển qua màn hình sẽ tạo thành 2 hình trụ trong đó hình dạng và kích thước có thể được xác định khi góc chiếu thay đổi. Từ đây, các câu

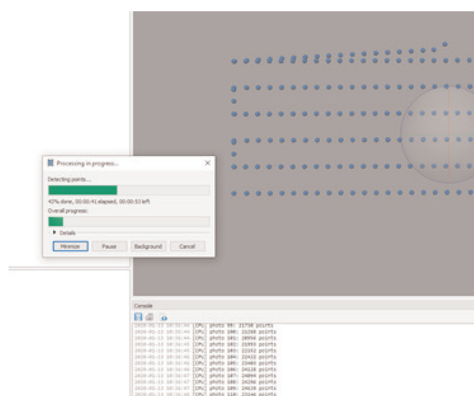
trúc từ chuyển động (SfM) có thể được tính toán.

Thuật toán SfM bao gồm 3 quá trình như hình dưới đây:



Hình 6: Các quá trình bên trong của SfM

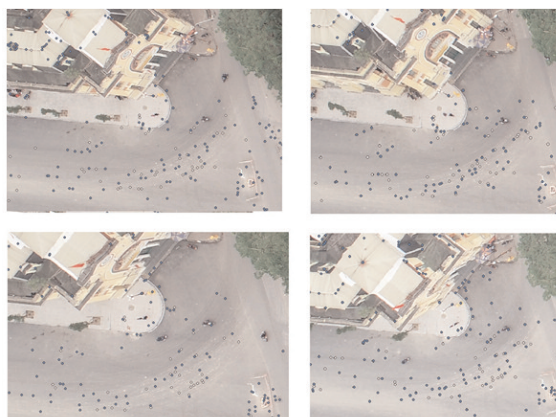
Bước đầu tiên của SfM là việc tìm kiếm các khu vực trùng nhau trên các tấm ảnh  $I = \{I_i | i=1 \dots N_I\}$  và xác định các phép chiếu hình trong các cặp ảnh. Kết quả của quá trình này là một tập hợp các cặp ảnh song trùng và phép chiếu hình cho mỗi điểm.



Hình 7: Quá trình phát hiện các điểm song trùng (Agisoft Photoscan)

Tại mỗi ảnh  $I_i$ , SfM phát hiện một tập  $F_i = \{(x_j, f_j) | j=1 \dots N_{F_j}\}$  của các đối tượng tại mỗi vị trí  $x_j$  thuộc  $R$  biểu diễn bởi tập  $f_j$ . Các đối tượng này sẽ có đặc tính hình học và độ xám thích hợp để SfM có thể tìm kiếm trong các hình ảnh khác

nhau tạo ra các điểm ảnh đặc trưng.



Hình 8: Tìm kiếm các điểm trùng nhau trên các bức ảnh (Agisoft Photoscan)

Các đối tượng dựa trên các điểm  $F_i$  trong các tấm ảnh sẽ được SfM tiến hành tìm kiếm và kết hợp. Các đối tượng có cấu trúc và độ sáng tương tự cho mỗi tấm ảnh... sẽ được SfM tìm kiếm trên toàn bộ khu vực ảnh trùng nhau. Các nhà khoa học đã đưa ra một số giải thuật để tăng cường hiệu suất kết hợp các đối tượng trên ảnh để thực hiện được công việc tìm kiếm đối với tập hợp ảnh lớn. Một tập hợp các điểm ảnh có khả năng trùng nhau  $C$  và các đối tượng trùng nhau  $M$  là kết quả của quá trình kết hợp đối tượng này.

Hiệu chỉnh các cặp ảnh trùng nhau  $C$  là giai đoạn thứ 3 trong kỹ thuật tìm kiếm ảnh song trùng. SfM sẽ hiệu chỉnh hình học các điểm ảnh trong các cặp ảnh trùng nhau bằng các phép chiếu hình. Phụ thuộc vào tính chất không gian của từng cặp ảnh, các hiệu chỉnh này sẽ thể hiện mối liên hệ hình học giữa chúng. Sự biến đổi xoay hoặc dịch chuyển thuần túy được miêu tả bằng các phép chiếu hình đồng nhất. Các phép chiếu đa cực được dùng cho các biến đổi phức tạp thông qua các ma trận chuẩn hóa và đa tiêu cực. Khi các biến đổi này được xây dựng đầy đủ cho các đối tượng trên ảnh, các đối tượng này được coi là đã hiệu chỉnh hình học. Tuy nhiên, việc xác định các phép chiếu này đối với ảnh thường có rất nhiều các nhiễu, do vậy, một số thuật toán như RANSAC hoặc GRIC hoặc

QDEGSAC sẽ được các chương trình SfM sử dụng để xử lý loại bỏ các nhiễu. Sau khi đã chạy hiệu chỉnh hình học, SfM cung cấp một đồ thị cảnh (scene graph) với ảnh là các đỉnh và các cặp ảnh được coi như các cạnh.

Trong số các thuật toán trên RANSAC được ứng dụng nhiều nhất, RANSAC được ứng dụng để ước tính quan hệ T với n tham số trên dữ liệu  $\{p\}$ . Theo đó, nó sẽ chia  $\{p\}$  thành cá giá trị trong và ngoài. Nó sẽ lựa chọn một tập m các đối tượng từ  $\{p\}$  và tính toán một ứng viên quan hệ Tc từ tập ngẫu nhiên này. Số lượng nhỏ nhất  $m = (n/r)$  của các đối tượng dùng để tính toán quan hệ phụ thuộc vào số lượng các hạn chế r được cung cấp bởi mỗi đối tượng và số lượng n các tham số của quan hệ T.

Bước thứ 2 của SfM là quá trình tăng cường tái cấu trúc. Đồ thị cảnh (scene graph) là đầu vào của quá trình tái cấu trúc. Một tập các ảnh đã được đăng ký P và tập hợp các điểm tái cấu trúc X là đầu ra của qua trình này.

SfM bắt đầu quá trình được gọi là khởi tạo mô hình bằng việc lựa chọn các cặp ảnh song trùng để tái xây dựng hình ảnh 3D là giai đoạn đầu tiên của quá trình tăng cường tái cấu trúc. Việc lựa chọn các cặp ảnh là việc làm hết sức quan trọng để đảm bảo chất lượng đầu ra. Hơn nữa, độ mịn, độ chính xác và hiệu suất của quá trình tái xây dựng phụ thuộc vào vị trí lựa chọn cặp ảnh trong quá trình tăng cường. Khởi tạo quá trình từ một tập hợp có mật độ cao các điểm song trùng sẽ giúp tăng cường chất lượng và độ chính xác của quá trình tái xây dựng hình ảnh 3D. Nếu khởi tạo quá trình từ một tập hợp có mật độ thấp sẽ dẫn đến sự kém hiệu quả và các bước tiếp theo sẽ gặp phải nhiều lỗi.

Đăng ký ảnh là giai đoạn 2 của quá trình tái cấu trúc. Quá trình này bắt đầu từ mô hình tái xây dựng ở bước khởi tạo. SfM sẽ đăng ký các ảnh vào mô hình bằng cách sử dụng các điểm tam giác song trùng để giải quyết vấn đề Hướng-điểm (Perspective-n-Point - PnP). PnP sẽ cho ra một tập hợp các vị trí P. Từ đây, tập P sẽ lần lượt được tăng cường bởi việc đăng ký liên tục các

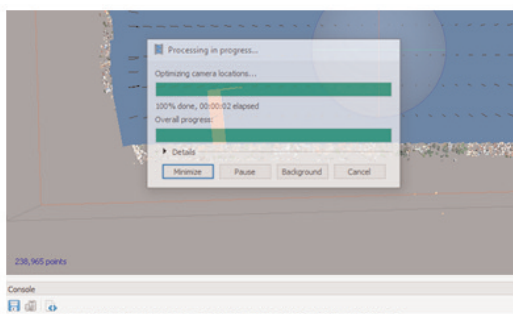
ảnh vào mô hình. Trong quá trình này, để loại bỏ các nhiễu thường gặp, các cặp ảnh được căn chỉnh và đưa vào giải thuật RANSAC để xử lý nhiễu.

Giai đoạn 3 của quá trình tái cấu trúc là bước bình sai. Trong quá trình bình sai tam giác mỗi tấm ảnh được đăng ký và tập hợp các điểm ảnh song trùng sẽ được mở rộng. Tập hợp các điểm song trùng sẽ lần lượt được tăng thêm khi có các ảnh được đưa vào mô hình. Bình sai tam giác là bước quan trọng trong SfM, nó giúp tăng cường sự ổn định của mô hình và cho phép mở rộng đăng ký cho các ảnh tiếp theo. Các nghiên cứu đã đưa ra nhiều phương pháp khác nhau [5][6] để thực hiện bình sai, hầu hết các phương pháp này có thể bị giới hạn bởi tính vững chắc của mô hình hoặc giới hạn tính toán trong SfM.

Giai đoạn cuối cùng của quá trình tăng cường tái cấu trúc là hiệu chỉnh toàn thể. Đăng ký và bình sai ảnh là các quá trình riêng rẽ mặc dù chúng đều cung cấp các điểm ảnh song trùng. Quá trình bình sai có thể lặp lại để mở rộng tập hợp các điểm ảnh của mô hình. Sau quá trình này, nếu không có thêm các hiệu chỉnh SfM sẽ đi tới giai đoạn không thể tái xử lý. Hiệu chỉnh toàn thể là một phương pháp hiệu chỉnh các tham số máy ảnh và các tham số điểm ảnh để giảm thiểu các sai số cho định tâm ảnh sử dụng hàm số Pi. Cấu trúc đặc biệt của các tham số sử dụng phương pháp của Schur [1] để hiệu chỉnh toàn thể. Thông thường, phương pháp này hoạt động rất tốt khi số lượng vị trí chụp nhỏ hơn số lượng điểm ảnh. Quá trình hiệu chỉnh toàn thể sẽ tiêu tốn rất nhiều thời gian và đòi hỏi công suất tính toán cao đối với các ảnh tập hợp không xấp xỉ. Quá trình này được đẩy nhanh do các tấm ảnh có sự sắp xếp và có cấu trúc, tiêu cự đồng nhất của ảnh chụp từ UAV.

Tìm kiếm các điểm 3D và các tham số máy ảnh nhằm giảm nhẹ các lỗi trong biến đổi phép chiếu hình là mục tiêu của quá trình hiệu chỉnh toàn bộ. Vấn đề tối ưu hóa này thường được công thức hóa như một dạng của số bình phương nhỏ nhất phi tuyến tính, khi lỗi là bình phương L2 biến trị vị trí giữa đối tượng quan sát và dựng

lên theo phép chiếu hình từ các điểm 3D theo các mặt phẳng ảnh của máy ảnh. Tuy nhiên, các dạng hàm số khác cũng được sử dụng như hàm của Huber không chỉ giới hạn trong việc sử dụng L2. Vấn đề có thể được tính theo dạng số bình phương nhỏ nhất có trọng số. Tuy nhiên, đối với hiệu chỉnh toàn khối (BA) đa phần sẽ dùng bình phương nhỏ nhất không tuyến tính.



Hình 9: Hiệu chỉnh toàn khối (Agisoft Photoscan)

Sau khi hiệu chỉnh toàn thể đám mây điểm, việc định vị đám mây điểm theo hệ thống tọa độ là điều hết sức quan trọng. Đám mây điểm được định vị dựa theo các mốc tiêu hay các điểm

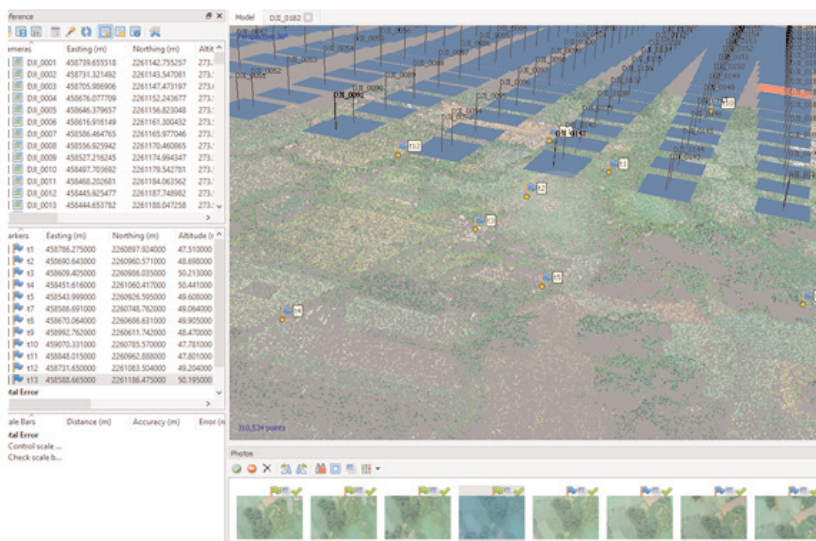
không chế ảnh (Ground Control Point – GCP) được chuẩn bị sẵn trong quá trình bay chụp. Việc định vị này có thể được thực hiện bởi các phần mềm miễn phí hoặc thông qua các modul định vị của các phần mềm thương mại như Agisoft PhotoScan. (Xem hình 10)

Mô hình điểm ảnh sẽ được tính toán tăng cường số lượng điểm để xây dựng được đám mây điểm tương ứng sau khi đã được định vị theo hệ thống tọa độ. (Xem hình 11)

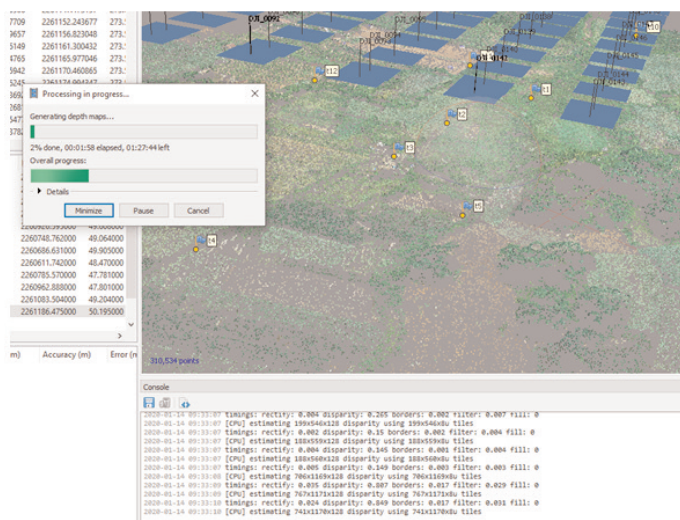
Một đám mây điểm có số lượng điểm lớn hơn rất nhiều (Đám mây điểm dày đặc - Dense Cloud) là kết quả của việc tái xây dựng mật độ điểm. Kết quả của quá trình SfM đó là đám mây điểm dày đặc (Dense Cloud). Đám mây điểm dày đặc này sẽ được sử dụng để tạo mô hình số độ cao (Digital Elevation Model - DEM) cho bước tiếp theo trong xử lý ảnh UAV.

### 3. Kết luận

Giải thuật biến đổi đối tượng phi tỷ lệ (SIFT) là một trong những giải thuật có giá trị ứng dụng cao trong lĩnh vực xử lý ảnh số. Giải thuật được đề xuất từ năm 1999 và đã được ứng dụng trong nhiều lĩnh vực trong đó nó đã làm thay đổi đáng kể trong lĩnh vực xử lý ảnh số. Phương pháp trích xuất các đặc trưng bất biến của SIFT được



Hình 10: Định vị đám mây điểm theo hệ tọa độ bằng các điểm không chế ảnh (Agisoft Photoscan)



Hình 11: Tái xây dựng mật độ điểm (Agisoft Photoscan)

thực hiện thông qua 4 bước là phát hiện điểm cực trị; định vị điểm đặc trưng; xác định hướng cho các điểm đặc trưng; mô tả đặc trưng thông qua các vector hướng. Việc trích xuất các điểm đặc trưng trong các ảnh lân cận sẽ giúp cho quá trình đăng ký và khớp ảnh đạt hiệu quả cao thông qua các điểm Keypoints. Trong giải thuật SIFT điểm đặc trưng của các đối tượng được trích xuất từ một tập hợp các hình ảnh tham chiếu và được lưu trữ trong cơ sở dữ liệu. Một đối tượng được nhận dạng trong một hình ảnh mới bằng cách so sánh riêng từng đặc trưng từ hình ảnh mới với cơ sở dữ liệu này và tìm các đặc trưng phù hợp với ứng viên dựa trên khoảng cách Euclide của các vector đặc trưng của chúng. Từ bộ đối sánh đầy đủ, các tập hợp con của các điểm chính đồng ý về đối tượng và vị trí, tỷ lệ và hướng của nó trong hình ảnh mới được xác định để lọc ra các kết quả khớp tốt.

Phương pháp tái tạo hình ảnh dựa vào chuyển động SfM áp dụng cho các tấm ảnh không sắp xếp là một kỹ thuật đã được phát triển mạnh mẽ, các hệ thống tái tạo hình học tự hiệu chỉnh là nền tảng đầu tiên của kỹ thuật SfM. Từ đây, các hệ thống tái tạo cấu trúc với quy mô lớn hơn ngày càng được phát triển để có thể xử lý hàng ngàn đến hàng triệu bức ảnh, cho đến nay, những hệ thống mới nhất có thể xử lý hàng trăm triệu bức

ảnh. Các phương pháp xử lý kỹ thuật SfM lần lượt được đưa ra như thuật toán tăng cường, thuật toán kế thừa, và các giải thuật toàn diện khác. Trong đó, giải thuật tăng cường là phương pháp xử lý thông dụng nhất đối với các ảnh chụp không sắp xếp. SfM được áp dụng cho các bức ảnh chụp từ thiết bị bay không người lái giúp tăng cường chất lượng đầu ra bao gồm đám mây điểm, ảnh ortho (UAV) và DSM (UAV). Những thành tựu từ các nghiên cứu này được áp dụng trong các phần mềm xử lý SfM thương mại và miễn phí đã và đang được sử dụng trong xử lý ảnh từ UAV.

Hiện nay, công tác xử lý dữ liệu UAV ở nước ta chủ yếu được sử dụng các phần mềm do nước ngoài phát triển, chưa có được các giải pháp phần mềm trong nước. Các phần mềm xử lý ảnh UAV này khá đa dạng bao gồm cả phần mềm mã nguồn mở và phần mềm thương mại. Các phần mềm này đều tích hợp các thuật toán SIFT, SfM,... cho phép gần như hoàn toàn tự động xử lý ảnh, xây dựng các sản phẩm bản đồ (mô hình số bề mặt, mô hình số độ cao, bản đồ trực ảnh, bản đồ 3D, video). Người sử dụng công nghệ này không nhất thiết phải có kiến thức quá sâu về công nghệ đo ảnh truyền thống.

Hơn thế nữa, công nghệ UAV với các tính



năng ưu việt của nó đã, đang và chắc chắn sẽ được ứng dụng rộng rãi hơn nữa vào các lĩnh vực khác nhau, đặc biệt là trong lĩnh vực đo đạc – bản đồ. Chính vì thế, nghiên cứu cơ sở khoa học của các thuật toán nói chung và giải thuật SIFT, SfM nói riêng sẽ giúp chúng ta hiểu được bản chất của vấn đề, trên cơ sở đó có thể xây dựng các chương trình và phần mềm tự động đăng ký và khớp ảnh. Công tác nghiên cứu các giải thuật SIFT kết hợp với một số các thuật toán khác như SfM sẽ giúp chúng ta có thể xây dựng các phần mềm tự động xử lý ảnh UAV mang thương hiệu Việt Nam trong tương lai.○

### **Tài liệu tham khảo**

[1]. S. Agarwal, N. Snavely, S. M. Seitz, and R. Szeliski, “Bundle Adjustment in the Large Computer Vision,” *Eccv*, vol. 6312, pp. 29–42, 2010.

[2]. Berni, J.A.J.; Zarco-Tejada, P.J.; Suarez, L.; Fereres, E., 2009, Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 47, pp. 722–738.

[3]. C. Caudek and N. Rubin, “Segmentation in structure from motion: Modeling and psychophysics,” *Vision Res.*, vol. 41, no. 21, pp. 2715–2732, 2001.

[4]. Kamarul Hawari Ghazali, “Feature

Extraction technique using SIFT keypoints descriptors“, *The International Conference on Electrical and Engineering and Informatics Institut technology Bandung, Indonesia*, 2007.

[5]. L. Kang, L. Wu, and Y. H. Yang, “Robust multi-view L2 triangulation via optimal inlier selection and 3D structure refinement,” *Pattern Recognit.*, vol. 47, no. 9, pp. 2974–2992, 2014.

[6]. H. Li, “A practical algorithm for  $L_{\infty}$  triangulation with outliers,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2007.

[7]. Lowe David, “Distinctive image features from scale-invariant keypoints“, *International Journal of Computer Vision*, 2004.

[8]. Lowe David, “Object recognition from local scale-invariant features“, *International Conference on Computer Vision, Corfu, Greece (September 1999)*, pp. 1150-1157.

[9]. Natan Micheletti, Jim H Chandler, Stuart N Lane, “Structure from Motion (SfM) Photogrammetry”, 2015.

[10]. Turner D., Luccieer A., Watson C, An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (SfM) Point Clouds, *Remote Sensing*, Vol. 4, pp. 1392-1410, 2012.○

### **Summary**

**UAV photogrammetry for topographical mapping based on Scale-Invariant Feature Transform and Structure from Motion**

*Nguyen Dai Dong*

*Department of Survey, Mapping and Geographic information Vietnam*

*Bui Ngoc Quy*

*Hanoi University of Mining and Geology*

The article about scientific basis of UAV photogrammetry for topographical mapping based on Scale-Invariant Feature Transform and Structure from Motion.○