

Thử nghiệm hệ thống định vị tích hợp thị giác lập thể và đo lường quán tính

An experiment in stereo vision aided inertial navigation system

Đặng Anh Tùng, Nguyễn Vĩnh Hảo, Vũ Ngọc Hải
Trường ĐH Bách Khoa TP.HCM, Trường ĐH Nguyễn Tất Thành
e-Mail: anhtunghcmut08@gmail.com, vinhhao@hcmut.edu.vn

Tóm tắt

Hệ thống định vị là một phần nghiên cứu cơ bản trong các lĩnh vực về giám sát và điều khiển vật thể chuyển động. Trong bài báo này nhóm đề xuất phương pháp tích hợp hệ thống định vị dùng thị giác lập thể và đo lường quán tính. Nhóm xây dựng các thuật toán độc lập cho từng hệ sau đó kết hợp lỏng dùng bộ lọc Kalman mở rộng để tích hợp lại. Giá trị ước lượng bao gồm vị trí và hướng của vật thể trong không gian 3D. Giải thuật đã được triển khai trên phần cứng thực tế và được lập trình theo mô hình nhiều tác vụ chạy song song để tận dụng tối đa khả năng tính toán của CPU nhằm tăng tốc giải thuật đáp ứng được các ứng dụng thời gian thực.

Từ khóa: INS, Stereo vision, localization.

Abstract

Localization system plays a fundamental role in the field of tracking and control moving objects. In this paper, we propose an integrated localization algorithm using stereo vision and inertial navigation system. We develop independent algorithm for each sub-system then fuse them into a loosely-couple scheme using extended Kalman filter. Estimation data included position, and orientation in 3D space. The method is embedded in industrial PC and use parallel programming model to boost the output data rate.

Keywords: INS, Stereo vision, localization.

Chữ viết tắt

INS	Inertial Navigation System
MEMS	Microelectromechanical systems
VO	Visual odometry

1. Giới thiệu chung hệ thống định vị tích hợp

Hệ thống dẫn đường quán tính INS là hệ thống dẫn đường sử dụng cảm biến gia tốc và cảm biến vận tốc góc quay để tính toán vị trí, và hướng của vật thể chuyển động mà không dùng thêm bất kỳ nguồn thông tin nào bên ngoài. Ban đầu hệ INS chỉ được sử dụng trong quân sự và hàng không. Để ứng dụng kỹ thuật này vào lĩnh vực robot di động, các nhà nghiên cứu sử dụng các loại cảm biến giá rẻ dùng công nghệ MEMS cho chất lượng thấp hơn. Điều này dẫn đến sai số định vị lớn do đó đặt ra yêu cầu có sự kết hợp với một loại

cảm biến khác để tăng độ chính xác. Công nghệ xử lý ảnh lập thể sử dụng một cặp ảnh đồng thời để ước lượng chuyển động. Với độ chính xác có thể đạt được như dùng GPS hoặc encoder nên nó được dùng tích hợp với bộ INS để định vị trong các môi trường mất tín hiệu GPS. Trong nghiên cứu này, nhóm đề xuất và kiểm nghiệm phương pháp tích hợp nhằm xây dựng một hệ thống dẫn đường sử dụng hai loại cảm biến trên.

Hệ thống INS hoạt động theo nguyên lý ước lượng gia tốc chuyển động của vật thể. Đầu tiên ta phải tính giá trị gia tốc chuyển động của robot dựa vào hiệu số giữa giá trị thu được từ cảm biến Accelerometer và hình chiếu của vector trọng trường Trái Đất lên robot. Sau đó ta tích phân hai lần để thu được quãng đường di chuyển. Phương pháp này có ưu điểm lớn là ta có thể phát triển thuật toán định vị độc lập với cấu trúc cơ khí của robot. Ta có thể chia sai số của phương pháp này làm hai phần. Đầu tiên là sai số hệ thống: do chất lượng cảm biến thấp hoặc lệch trục tương đối giữa IMU-mặt đất, IMU-vật thể; Và sai số ngoài hệ thống: do giải thuật ước lượng IMU chưa tối ưu hay vật thể di chuyển chậm dẫn đến giá trị nhiễu chiếm phần lớn trong giá trị gia tốc. Sai số hệ thống sẽ được cải thiện nhờ vào quá trình cân chỉnh hệ thống, và tăng chất lượng của cảm biến [1]. Còn sai số ngoài hệ thống có thể được cải thiện nhờ tối ưu giải thuật như sử dụng bộ lọc Kalman [2,3,4,5]. Ngoài ra có thể tích hợp thêm cảm biến Magnetometer để tăng độ chính xác cho bộ IMU. Tuy nhiên việc tích phân hai lần trên giá trị gia tốc vẫn sẽ gây ra sai số tích lũy lớn dần theo thời gian và vẫn chưa thể áp dụng vào định vị [6]. Do đó đặt ra yêu cầu phải tích hợp thêm cảm biến để bù lại sai số này. Đó cũng là lý do cộng đồng nghiên cứu về hệ định vị phải kết hợp hệ INS với encoder ở bánh xe, GPS hoặc/và công nghệ xử lý ảnh để tăng độ chính xác cho bộ định vị.

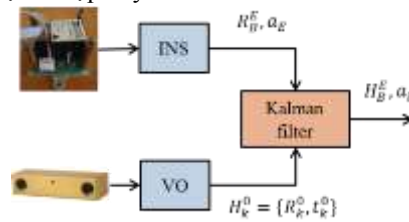
Phương pháp định vị dùng công nghệ xử lý ảnh từ camera gọi là Visual Odometry (VO). Phương pháp này đang được dùng rộng rãi vì giá thành hợp lý đồng thời khả năng mở rộng các tính năng khác ngoài định vị của camera là rất lớn. Ví dụ như định vị kết hợp lập bản đồ SLAM, tránh vật cản, dò biển báo giao thông [7,8]... Hệ thống định vị dùng công nghệ xử lý ảnh có nhiều ưu điểm vượt trội so với các công nghệ khác: sử dụng được trong môi trường trong nhà lẫn ngoài trời; cách ly được ảnh hưởng bởi hiện tượng trượt của

robot; có thể tích hợp mở rộng thêm nhiều chức năng khác như dò để tránh vật cản, dò và bám, xây dựng lại bản đồ 3D quang đường đã di chuyển (3D SLAM). Tất nhiên bên cạnh đó hệ thống định vị dùng công nghệ xử lý ảnh vẫn có các nhược điểm đồng thời cũng là các thách thức đặt ra cho việc nghiên cứu như: hệ thống định vị tương đối nên phụ thuộc rất nhiều vào các căn chỉnh ban đầu như vị trí, góc nghiêng, các thông số nội của camera. Hơn nữa sai số hệ thống sẽ tích lũy theo thời gian nên đặt ra yêu cầu phải có thêm phần hiệu chỉnh lại vị trí sau một thời gian dài di chuyển; nhạy với nhiễu từ các nguồn sáng: là vấn đề mà tất cả các hệ thống xử lý ảnh phải đối mặt; phụ thuộc vào mức độ giàu có của môi trường về thông tin: môi trường càng nhiều đặc trưng như cây cối, đồ vật... thì càng phù hợp sử dụng công nghệ xử lý ảnh [7,8].

Xử lý ảnh lập thể (Stereo-Vision) là một nhánh trong công nghệ định vị dùng xử lý ảnh VO. Đây là công nghệ tính toán vị trí và hướng chuyển động của camera (hay vật thể được gắn camera) sử dụng các cặp ảnh liên tiếp nhau. Nó là một phần trong hướng nghiên cứu lớn hơn là tái dựng lại môi trường 3D từ chuyển động camera (Structure-from-Motion). Ta cũng có thể sử dụng camera đơn tuy nhiên nhược điểm chính của camera đơn là chỉ ước lượng được chuyển động thành các giá trị tương đối. Nhiều nghiên cứu về VO đã chỉ ra rằng trong các điều kiện cụ thể sai số đạt được của hệ có thể dưới 2% quang đường di chuyển [15, 16]. Độ chính xác này chứng minh tính khả thi của VO khi dùng cho các hệ thống dẫn đường. Nghiên cứu của Geiger [16] đã trình bày một giải thuật hiệu quả để ước lượng chuyển động từ hai cặp ảnh liên tiếp nhau trong thời gian thực. Sai số đạt được từ phương pháp của Geiger [16] được nhóm phát triển và kiểm nghiệm khi chạy theo vòng tròn có sai số lớn nhất là 0.36m trên tổng quang đường 15m (2.5%). Từ kết quả thử nghiệm trên nhóm quyết định sử dụng giải thuật VO này cho stereo-camera. Và ngõ ra của hệ VO sẽ là đầu vào cho mô hình đo lường của bộ tích hợp.

Tóm lại hệ INS và VO riêng lẻ có những ưu, khuyết điểm như: hệ INS nhạy với nhiễu khi đối tượng hoạt động ở chuyển động chậm nhưng cho ra kết quả tốt hơn khi ở chuyển động nhanh; hệ INS cho tốc độ cập nhật rất cao vì giải thuật tính toán tinh gọn; hệ VO có thể ước lượng tốt trong trường hợp chuyển động chậm, tốc độ càng nhanh thì càng dễ mất các đặc trưng do đó sẽ định vị không hiệu quả. Đồng thời tốc độ cập nhật hệ VO luôn thấp vì các giải thuật xử lý ảnh chiếm nhiều thời gian tính toán. Như vậy hệ INS với khả năng cho tần số cập nhật cao, cung cấp ma trận xoay chính xác và gia tốc ngoài trong trường hợp robot di chuyển tốc độ cao sẽ là giải pháp hỗ trợ tốt cho camera khi mà hầu hết các giải thuật xử lý ảnh cho tần số cập nhật thấp sẽ tạo ra khoảng trễ không phù hợp với các bài toán điều khiển. Vì vậy sự kết hợp hai hệ thống này là điều cần thiết để bổ khuyết các nhược điểm của từng hệ riêng lẻ nhằm đạt độ chính xác cao hơn.

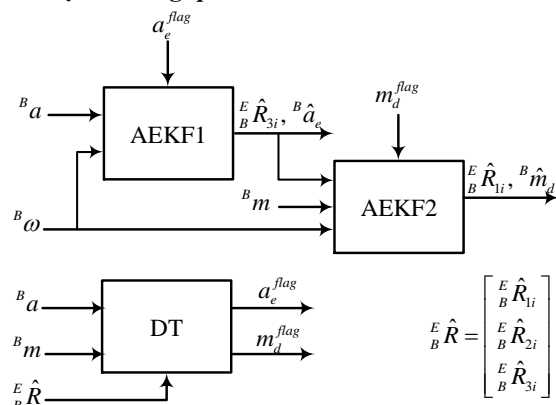
Có nhiều phương pháp tiếp cận để giải quyết bài toán tích hợp hai hệ thống này nhưng nói chung chia làm hai loại đó là: cấu trúc kết hợp lỏng (loosely-couple), và cấu trúc kết hợp chặt (tightly-couple). Trong cấu trúc kết hợp chặt các bước đầu quá trình xử lý ảnh được kết hợp luôn vào quá trình ước lượng. Một lựa chọn phổ biến là hỗ trợ dò các điểm đặc trưng [9, 10]. Hoặc là dùng thông tin từ IMU để đưa ra các giới hạn về ebipolar được áp dụng trong một số ứng dụng thời gian thực [11,12]. Khác với cấu trúc trên thì kết hợp lỏng phát triển riêng biệt hai bộ INS và Vision, chạy ở các tần số khác nhau và đưa giá trị đã xử lý về một bộ lọc trung tâm để kết hợp. Giá trị từ xử lý ảnh làm đầu vào cho mô hình đo lường của bộ tích hợp. Thông tin có thể lấy là vector dịch chuyển và ma trận xoay của hai frame ảnh liên tiếp nhau [13,14]. Cấu trúc lỏng có ưu điểm nổi bật là đơn giản hơn và cho phép tách riêng giải thuật xử lý ảnh với giải thuật tích hợp. Mặt khác nhiều giải thuật VO đã đạt đến độ chính xác rất cao nên nhóm có thể tận dụng lại để nhanh chóng xây dựng hoàn thiện một bộ tích hợp định vị. Do đó trong nghiên cứu này nhóm quyết định sẽ dùng cấu trúc kết hợp lỏng để tiến hành xây dựng và kiểm nghiệm bộ định vị kết hợp này.



Hình 1. Mô hình tích hợp lỏng hệ thống định vị
 Hệ thống tích hợp này có thể chạy độc lập trong trường hợp định vị cục bộ hoặc là hỗ trợ cho các bộ thu GPS chính xác cao trong trường hợp mất tín hiệu vệ tinh.

2. Phương pháp thực hiện

2.1 Hệ đo lường quán tính INS



Hình 2. Sơ đồ khối bộ ước lượng EKF thích nghi hai tầng

Ở đây để xây dựng giải thuật cho riêng bộ INS nhóm sử dụng giải thuật bộ lọc Kalman thích nghi mở rộng dùng phương pháp DCM đã phát triển ở [20]. Giải thuật hệ INS bao gồm 2 bộ lọc thích nghi mở rộng Kalman nối tiếp nhau. Bộ AEKF1 sẽ ước lượng ba giá

trị hàng thứ ba trong ma trận DCM, và gia tốc ngoài trên 3 trục. Bộ AEKF2 sẽ ước lượng ba giá trị hàng thứ nhất trong ma trận DCM và nhiễu từ trường trên ba trục. Hàng còn lại sẽ được suy ra dựa vào tính chất trực giao trong ma trận DCM.

Yêu cầu của các phương pháp này là phải có điều kiện cần để “kích hoạt” bộ thích nghi. Ở điều kiện thường (gia tốc ngoài nhỏ hoặc không có nhiễu từ trường ngoài) ta sẽ “ngắt” bộ này ra khỏi mô hình (bằng cách thiết lập giá trị các biến là 0); khi có điều kiện cần ta sẽ “đóng” các biến này vào mô hình. Thông thường ta sẽ dùng giá trị độ lớn là hiệu số tổng gia tốc đo được so với tổng gia tốc khi đứng yên và hiệu số tổng từ trên ba trục so với tổng từ cố định của Trái Đất tại vị trí đang xét. Nếu giá trị hiệu này vượt một ngưỡng cho trước thì sẽ “kích hoạt” bộ thích nghi. Giá trị ngưỡng nếu quá lớn sẽ dẫn đến hệ thống chậm thích nghi với điều kiện ngoài, còn nếu quá nhỏ sẽ dẫn đến hệ thống nhạy với nhiễu.

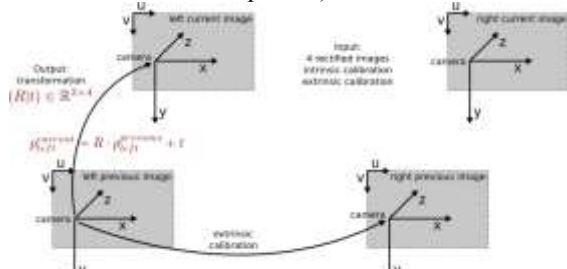
Kết quả ngõ ra thử nghiệm trên bàn xoay dùng encoder cho độ chính xác rất cao khoảng 2° trên cả ba góc Euler [20].

2.2 Hệ thị giác lập thể stereo-vision

Trong bài báo này nhóm sử dụng giải thuật được trình bày ở [16] để tiến hành thực nghiệm đánh giá kết quả của hệ định vị dùng stereo-camera. Phương pháp để ước lượng chuyển động dùng stereo-camera (hay còn gọi là visual odometry) có thể tóm tắt qua các bước như sau:

Bước 1: Dò điểm đặc trưng

Ngõ vào cho giải thuật visual odometry là những điểm đặc trưng dò được giữa bốn ảnh (2 ảnh trái phải và 2 frame ảnh liên tiếp nhau) như Hình 3.



Hình 3. Mô hình ước lượng chuyển động của Stereo camera

Để tìm vị trí đặc trưng ổn định, bài báo sử dụng bộ lọc ra các đặc trưng góc (corner) & đốm (blob). Sau đó áp dụng thuật toán non-maximum- and non-minimum-suppression trên ảnh và kết quả các đặc trưng phân loại vào 4 class: blob max, blob min, corner max, corner min.

Từ hai điểm đặc trưng so sánh khối cửa sổ 11x11 the chiều ngang và chiều dọc bộ lọc Sobel sử dụng SAD (sum of absolute differences). Giải thuật ước lượng chuyển động bắt đầu bằng matching những đặc trưng bên ảnh trái hiện tại với ảnh trái trước đó trong vùng cửa sổ $M \times M$, sau đó tìm tới ảnh bên phải trước đó, sau đó tới ảnh phải hiện tại, rồi lại ảnh trái hiện tại. Việc dò tìm theo vòng tròn này đảm bảo lấy được đặc trưng ổn định.

Bước 2: Ước lượng chuyển động

Từ các đặc trưng, ta tính chuyển động của camera dựa vào tối thiểu hóa hàm sai số

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & c_u \\ 0 & f & c_v \\ 0 & 0 & 1 \end{pmatrix} \left[(\mathbf{R}(\mathbf{r}) \quad \mathbf{t}) \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} - \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} \right]$$

Với:

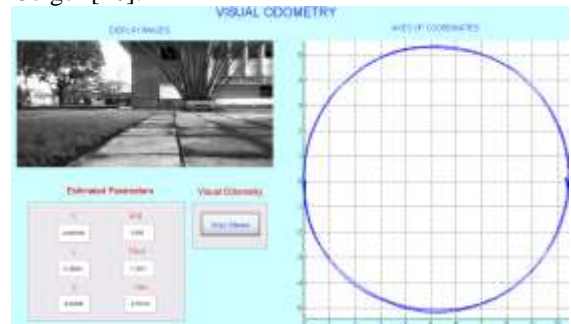
- pixel trên hệ trục ảnh $(u, v, 1)$
- tiêu cự camera f
- quang tâm (c_u, c_v)
- ma trận xoay R
- vector dịch $\mathbf{t} = (t_x, t_y, t_z)$
- điểm trong mặt phẳng 3D (x, y, z)
- baseline của stereo s

Sử dụng phép lặp Gauss-Newton tối thiểu hóa hàm mục tiêu là tổng sai số của hình chiếu các điểm 3D xuống lần lượt các ảnh trái và phải để tìm ma trận xoay và vector dịch chuyển ở hai thời điểm bắt ảnh liên tiếp nhau.

$$\sum_{i=1}^N \left\| \mathbf{x}_i^{(l)} - \pi^{(l)}(\mathbf{X}_i; \mathbf{r}, \mathbf{t}) \right\|^2 + \left\| \mathbf{x}_i^{(r)} - \pi^{(r)}(\mathbf{X}_i; \mathbf{r}, \mathbf{t}) \right\|^2$$

Với p là hàm chiếu vector điểm 3D xuống mặt phẳng ảnh 2D.

Sai số đạt được kiểm nghiệm từ phương pháp của Geiger [16]:



Hình 4. Thử nghiệm chạy vòng tròn theo phương pháp của Geiger [16]. Sai số lớn nhất là 0.36m trên tổng quãng đường 15m (2.5%).

Từ kết quả thử nghiệm trên nhóm quyết định sử dụng giải thuật VO này cho stereo-camera. Và ngõ ra của hệ VO sẽ là đầu vào cho mô hình đo lường của bộ tích hợp.

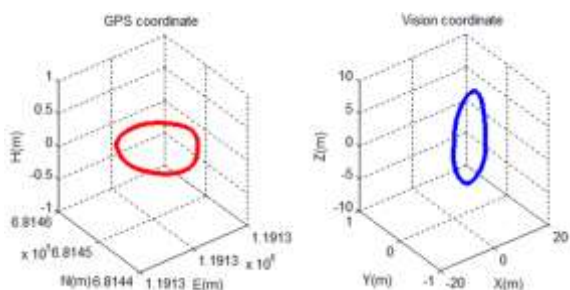
3. Tích hợp hệ thống

3.1 Cân chỉnh hệ thống

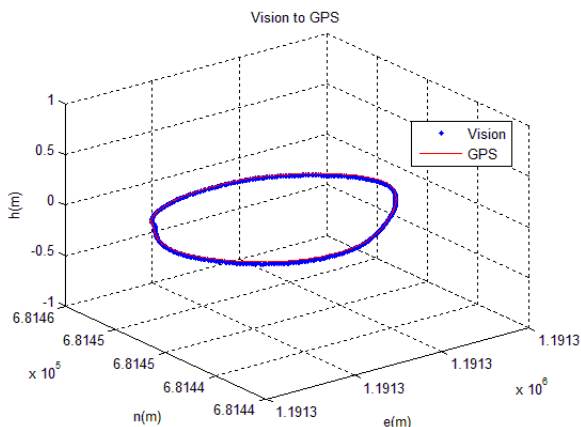
Trong nghiên cứu này để tiến hành thực nghiệm nhóm phải sử dụng ba hệ trục trong các hệ thống bao gồm:

- INS: {I} dùng NED
 - GPS: {G} dùng hệ ENU
 - Stereo Vision: {V} dùng hệ cục bộ Decac XYZ
- Do đó để tích hợp hay đánh giá các hệ thống định vị khác nhau này ta phải cân chỉnh bằng cách tìm ma trận chuyển đổi giữa các hệ trục. Ma trận chuyển đổi đầy đủ bao gồm một ma trận xoay và vector tịnh tiến.

Việc cân chỉnh hệ GPS/INS/Stereo-camera có hai ý nghĩa. Ý nghĩa đầu tiên là nhóm sử dụng tín hiệu GPS là tín hiệu chuẩn (ground-truth data) để so sánh và đánh giá hệ INS/Stereo-camera. Ý nghĩa thứ hai là đánh giá tính khả thi của việc kết hợp hệ stereo-camera trong việc hỗ trợ hệ GPS khi bộ thu mất tín hiệu. Điều này thường xảy ra khi đối tượng hoạt động ở các môi trường đô thị hoặc vùng cây cối che khuất. Phương pháp cân chỉnh hệ INS/Stereo sử dụng công cụ cân chỉnh hay dùng của cộng đồng xử lý ảnh do Bouguet phát triển [17]. Tiếp theo ta chỉ cần tìm ma trận chuyển giữa hệ INS/Stereo-camera và GPS. Nhiều phương pháp khác nhau cân chỉnh hai hệ này dùng giải thuật đệ quy và không cần đệ quy, trong bài báo này nhóm chọn giải pháp dùng phương pháp không đệ quy cụ thể dùng SVD (singular value decomposition) [21]. Ngõ vào của thuật toán sẽ là hai bộ điểm từ hai hệ trục, ngõ ra sẽ là ma trận chuyển đổi giữa hai hệ. Kết quả cân chỉnh hai hệ thống được trình bày trong Hình 5 và Hình 6.



Hình 5. Hai đường di chuyển trong hai hệ trục GPS và xử lý ảnh trước khi cân chỉnh



Hình 6. Hai đường di chuyển trong hai hệ trục GPS và xử lý ảnh sau khi cân chỉnh về chung hệ Vision

3.2 Giải thuật tích hợp

Mô hình tổng quát của hệ thống tích hợp bất đồng bộ:

$$x = \begin{bmatrix} p \\ v \\ a \end{bmatrix}^T$$

Với:

- $p = [p_x, p_y, p_z]^T$: vị trí của camera trong không gian, trong trường hợp này hệ quy chiếu là mặt đất với gốc 0 là điểm bắt đầu của hệ thống.

- $v = [v_x, v_y, v_z]^T$: vận tốc chuyển động theo ba trục.

- $a = [a_{ex}, a_{ey}, a_{ez}]^T$: gia tốc chuyển động (ta kí hiệu thêm chữ 'e' – external để phân biệt giá trị này với giá trị gia tốc từ cảm biến Accelerometer).

Vì INS và Stereo camera có tần số cập nhật khác nhau và để tăng tốc độ cập nhật lên tối đa bộ ước lượng sẽ chạy ở tần số 100Hz trùng với tần số cập nhật của INS. Biến đo lường cập nhật bao gồm đầy đủ giá trị vị trí tính từ Camera và gia tốc ngoài tính từ INS:

$$y = [p_{cam}, a_{imu}]^T$$

Nhưng ở tần số 100Hz hệ thống có thể có hai mode cập nhật khác nhau tùy theo giá trị có sẵn khi đó của các cảm biến. Các khả năng có thể xảy ra là:

$$y_0 = [p_{cam}, a_{imu}]^T$$

$$y_1 = [a_{imu}]^T$$

Mô hình quá trình: (Process model - ở đây gia tốc ngoài được mô hình thành bộ lọc thông thấp bậc 1)

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \quad (1)$$

Từ (1) ta rút ra mô hình quá trình như sau:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = f x \quad (2)$$

Mô hình đo lường: (Measurement model)

$$\begin{bmatrix} p_{cam} \\ a_{imu} \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} x \quad (3)$$

Ta rút ra mô hình đo lường ở các điều kiện khác nhau:

$$y_0 = [p_{cam}, a_{imu}]^T \quad y_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = H_0 x \quad (4)$$

$$y_1 = [a_{imu}]^T \quad y_1 = [0 \ 0 \ 0] x = H_1 x \quad (5)$$

Từ mô hình trên ta tiến hành rời rạc hóa để áp dụng bộ EKF [19] với phương pháp xấp xỉ bậc nhất

$$\begin{aligned} \hat{x}_k &= (I + fT)x_{k-1} = f_k x_{k-1} \\ \hat{y}_k &= Hx_k = H_k x_k \end{aligned} \quad (6)$$

Ma trận covariance của nhiễu quá trình: được suy ra từ mô hình quá trình. Tuy nhiên ở đây ta giả sử các thành phần covariance trên ba trục x, y, z của gia tốc ngoài là độc lập nhau và là variance của chính nó.

$$Q_k = \begin{bmatrix} s_{ea}^2 T^2 & \frac{1}{2} s_{ea}^2 T^3 & \frac{1}{2} s_{ea}^2 T^2 \\ \frac{1}{2} s_{ea}^2 T^3 & s_{ea} T & s_{ea} T^2 \\ \frac{1}{2} s_{ea}^2 T^2 & s_{ea} T^2 & s_{ea} \end{bmatrix} \quad (7)$$

Ma trận covariance của nhiễu đo lường: giả sử chỉ có thành phần covariance trên đường chéo.

$$y_0 = [P_{cam}, a_{imu}]^T \quad R_{0,k} = \begin{bmatrix} \hat{e}^{cam} & 0 \\ 0 & s_{imu} \end{bmatrix} \quad (8)$$

$$y_1 = [a_{imu}]^T \quad R_{1,k} = [s_{imu}] \quad (9)$$

4. Thí nghiệm

4.1 Mô hình phần cứng

Nhóm sử dụng bộ GPS-RTK độ chính xác cấp cm là giá trị chuẩn để so sánh với kết quả bộ tích hợp.

Một bộ kit thí nghiệm được xây dựng bao gồm:

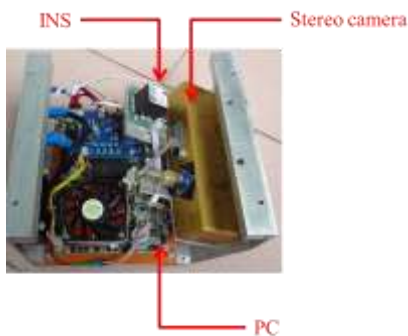
- Board máy tính công nghiệp NANO-HM650: Intel Core i5, ổ cứng SSD 120GB được dùng để giúp máy tính chạy ổn định vì hoạt động ngoài trời và dưới các điều kiện rung lắc lớn, sử dụng hệ điều hành Window XP SP3.

- Stereo camera Point Grey Bumblebee2: các cặp ảnh xám được đọc vào máy tính thông qua chuẩn Firewire IEEE-1394.

- IMU Analog Devices ADIS16407: dữ liệu của các cảm biến gia tốc, vận tốc góc, và từ trường được lấy thông qua chuẩn SPI với tần số 100Hz, sử dụng một board vi điều khiển để đọc và truyền dữ liệu này lên máy tính thông qua RS232.

- GPS RTK South độ chính xác cấp cm: dữ liệu đọc vào máy tính thông qua RS232 tần số 10Hz được đồng bộ với dữ liệu INS.

Để thu giá trị chuẩn GPS chính xác ta phải di chuyển hệ ngoài trời. Với mục tiêu áp dụng cho các hệ robot trong nông nghiệp hoặc xe tự hành ngoài trời nên nhóm đặt camera cao khoảng 2m so với mặt đất trên một xe kéo bằng tay. Và để hạn chế phần nửa ảnh trên chỉ chụp bầu trời khi đặt camera cao thì phần cứng có hiệu chỉnh camera nghiêng xuống một góc 30°.



Hình 6. Phần cứng xây dựng hoàn chỉnh

4.2 Mô hình phần mềm

Để tối ưu thời gian tính toán và tăng tốc chương trình tích hợp, phần mềm được viết hoàn toàn dựa trên ngôn ngữ C++. Nhằm tăng tính độc lập và dễ mở rộng phần mềm tạo ra 3 class độc lập và mỗi class chứa một thread lập vô cùng. Tất cả các thread chạy song song nhau để quản lý các tác vụ khác nhau bao gồm:

Vision thread:

- Thu nhận cặp ảnh từ Stereo camera dùng thư viện cung cấp bởi nhà sản xuất camera. Thư viện này cung cấp các API để truy cập tới camera.

- Chạy giải thuật xử lý ảnh để ước lượng ra ma trận chuyển vị hai frame ảnh liên tiếp nhau. Ma trận chuyển vị này bao gồm một ma trận xoay và một vector dịch chuyển.
- Giá trị này được chuyển sang cho Serial thread để tích hợp vào bộ ước lượng.

Serial thread:

- Nhận dữ liệu từ cổng nối tiếp RS232 bao gồm dữ liệu INS, GPS.
- Tách chuỗi và tính toán các giá trị thực từ các chuỗi này.
- Dùng dữ liệu INS làm sự kiện trigger cho giải thuật tích hợp.

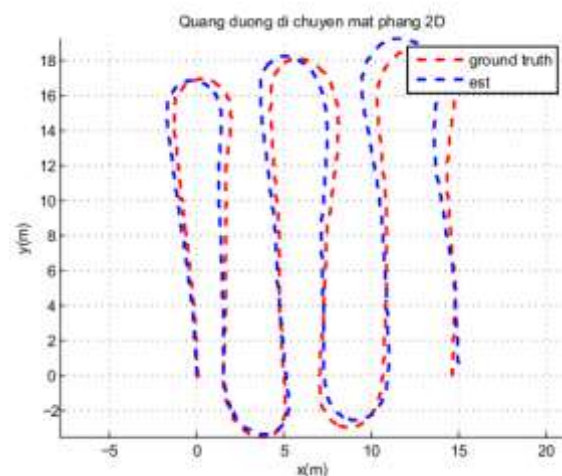
Main thread:

- Đóng vai trò trung tâm khởi tạo các thread.
- Kiểm tra dữ liệu cập nhật mới từ Vision thread để chuyển qua Serial thread.
- Kiểm tra trạng thái các chân IO để bật còi.

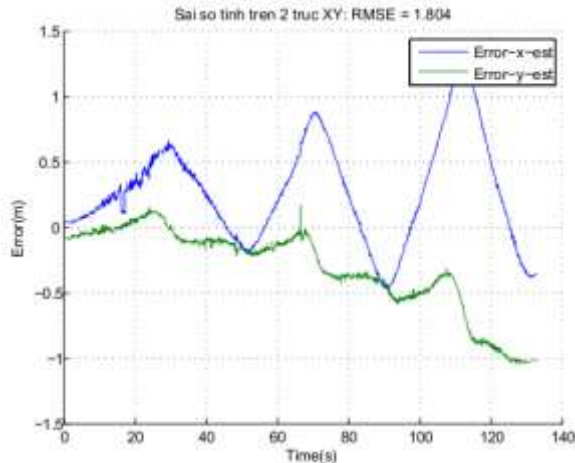
4.3 Kết quả thực nghiệm

Mục tiêu của nghiên cứu nhằm xây dựng hệ tích hợp có thể chạy độc lập hoặc hỗ trợ cho hệ GPS trong các trường hợp mất tín hiệu vệ tinh. Do đó trước tiên nhóm tiến hành thí nghiệm kéo hệ thống chạy lên xuống trong sân trường để kiểm tra độ chính xác. Sau đó mô phỏng quá trình hỗ trợ cho hệ GPS khi mất tín hiệu vệ tinh bằng cách chuyển ngõ ra từ GPS sang ngõ ra cập nhật từ hệ INS/Vision. Hai ngõ ra này được cân chỉnh hệ trục dùng phương pháp cân chỉnh đã mô tả ở phần 3. Vì hệ thống sử dụng công nghệ xử lý ảnh do đó môi trường thử nghiệm hệ thống tích hợp này phải được chọn lọc chỉ sử dụng ở môi trường phải có đặc trưng nhiều. Ngoài ra nếu thử nghiệm ở các môi trường ít đặc trưng hoặc camera bị chói sáng thì hệ thống không thể phát huy hiệu quả của mình do bản chất vật lý của hệ.

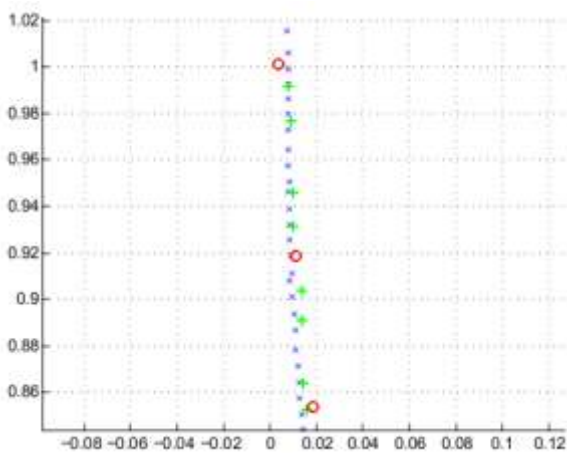
Kết quả 1: (Hình 7-8-9) Di chuyển hệ thống lên xuống trên đường phẳng với tổng quãng đường di chuyển hơn 100m. Sai số tối đa trên hai trục là 1.5m. Sai số RMS là 1.8m. Ngõ ra đạt độ chính xác khá cao có thể dùng được trong các ứng dụng giám sát.



Hình 7. Quãng đường di chuyển



Hình 8. Sai số trên 2 trục



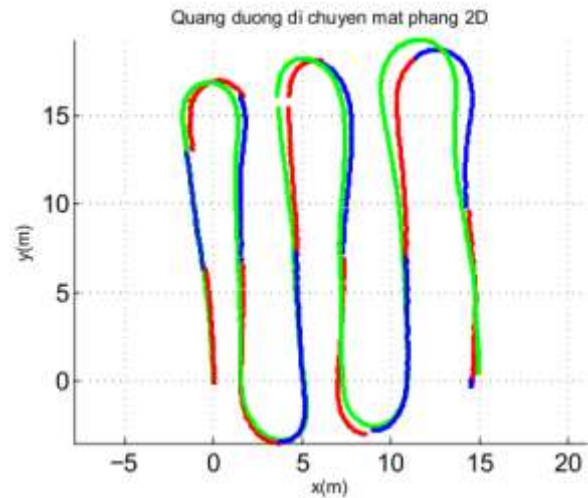
Hình 9. Phóng to các điểm ước lượng. Màu đỏ là giá trị cập nhật GPS 10Hz, màu xanh lá là ngõ ra giải thuật xử lý ảnh từ 20-40Hz, màu xanh dương là điểm ước lượng 100Hz.

Kết quả 2: (Hình 10) Mô phỏng quá trình mất GPS theo chu kỳ 10s và chuyển sang dùng giá trị ước lượng từ hệ INS/Stereo-camera. Đường màu đỏ là giá trị chuẩn từ GPS-RTK. Đường màu xanh lá là kết quả từ ngõ ra hệ tích hợp. Đường màu xanh da trời là kết quả vị trí khi mất GPS và sử dụng hệ INS/Stereo-camera có cân chỉnh lại mỗi khi mất GPS dùng phương pháp ở Phần 3. Kết quả cho thấy hệ thống đã tích hợp hoàn toàn có thể hỗ trợ bộ GPS khi mất tín hiệu vệ tinh.

5. Kết luận

Nhóm nghiên cứu đã trình bày một giải thuật hiệu quả để tích hợp hệ thống xử lý ảnh lập thể vào hệ thống đo lường quán tính nhằm nâng cao chất lượng của hệ thống định vị. Nhóm đã phát triển mô hình tích hợp dùng bộ lọc EKF theo cấu trúc kết hợp lồng biểu diễn đầy đủ vị trí cũng như hướng của vật thể trong không gian. Nhóm cũng đã trình bày mô hình phần cứng và phần mềm đã xây dựng nhằm chạy hệ thống trong thời gian thực. Trong tương lai nhóm sẽ tiến hành

nhiều thử nghiệm khác để đánh giá tính ổn định của hệ thống.



Hình 10. Di chuyển và mất GPS theo chu kỳ 10s

Lời cảm ơn

Dự án này được hỗ trợ bởi Bộ Khoa học và Công nghệ thông qua đề tài mã số KC.03.15/11-15 và Trường Đại học Bách Khoa, Đại học Quốc Gia TP.HCM trong khuôn khổ đề tài mã số T-ĐĐT-2015-13, C2014-20-03.

Tài liệu tham khảo

- [1] Benjamin Peter, "Development of an Automatic IMU Calibration System" *Master Thesis*, ETHZ, 2011.
- [2] Bachmann, E.R. *et al*, "Orientation Tracking for Humans and Robots Using Inertial Sensors", *In 1999 International Symposium on Computational Intelligence in Robotics & Automation*, 1999, pp.187-194.
- [3] Marins, J.L., Yun, X., Bachmann, E.R., McGhee, R.B., Zyda, M.J., "An Extended Kalman Filter for Quaternion-Based Orientation Estimation Using MARG Sensors", *In Proceedings of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2001, vol.4, pp. 2003-2011.
- [4] Yun, X., Bachmann, E.R., "Design, Implementation, and Experimental Results of a Quaternion-Based Kalman Filter for Human Body Motion Tracking", *In IEEE Transactions on Robotics*, 2006, vol. 22, pp. 1216-1227.
- [5] Daniel Roetenberg, *Inertial and Magnetic Sensing of Human Motion*, PhD Thesis, 2006, University of Twente.
- [6] Woodman, Oliver J. "An introduction to inertial navigation" *Technical Report*, 2007.
- [7] D. Scaramuzza, F. Fraundorfer, "Visual Odometry: Part I - The First 30 Years and Fundamentals", *IEEE Robotics and Automation Magazine*, vol.18(4), 2011, pp. 80-92.

- [8] F. Fraundorfer, D. Scaramuzza, "Visual Odometry: Part II - Matching, Robustness and Applications", *IEEE Robotics and Automation Magazine*, vol.19(2), 2012, pp. 78–90.
- [9] M. Veth and J. Raquet, "Fusion of low-cost imaging and inertial sensors for navigation," *Air Force Institute of Technology*, Tech. Rep., 2007.
- [10] J. Jurado, K. Fisher, and M. Veth, "Inertial and imaging sensor fusion for image-aided navigation with affine distortion prediction," in *Position Location and Navigation Symposium (PLANS)*, 2012, pp. 518–526.
- [11] J.-O. Nilsson, D. Zachariah, M. Jansson, and P. Handel, "Realtime implementation of visual-aided inertial navigation using epipolar constraints," in *Position Location and Navigation Symposium (PLANS)*, 2012, pp. 711–718.
- [12] D. D. Diel, P. DeBitetto, and S. Teller, "Epipolar constraints for vision-aided inertial navigation," in *IEEE Workshop on Motion and Video Computing*, 2005, pp. 221–228.
- [13] J.-P. Tardif, M. D. George, M. Laverne, A. Kelly, and A. Stentz, "A new approach to vision-aided inertial navigation." in *IROS. IEEE*, 2010, pp. 4161–4168.
- [14] S. Roumeliotis, A. Johnson, and J. Montgomery, "Augmenting inertial navigation with image-based motion estimation" in *Robotics and Automation ICRA*, vol. 4, 2002, pp. 4326 – 4333.
- [15] M. Maimone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers: Field reports," *J. Field Robot.*, vol. 24, no. 3, pp. 169–186, 2007.
- [16] A. Geiger, J. Ziegler, and C. Stiller. "Stereoscan: Dense 3d reconstruction in real-time." *Intelligent Vehicles Symposium (IV)*, 2011 IEEE.
- [17] J. Bouguet, "Visual methods for three-dimensional modeling," Ph.D. dissertation, California Institute of Technology, Pasadena, California, 1999.
- [18] Lobo, Jorge, and Jorge Dias. "Relative pose calibration between visual and inertial sensors." *The International Journal of Robotics Research* 26.6 (2007): 561-575.
- [19] Greg Welch and Gary Bishop (2006). An Introduction to the Kalman Filter.
- [20] Anh-Tung Dang, Vinh-Hao Nguyen "DCM-Based Orientation Estimation Using Cascade of Two Adaptive Extended Kalman Filters" *2013 International Conference on Control, Automation and Information Sciences*, pp. 152-157, IEEE, 2013.
- [21] Arun, K. Somani, Thomas S. Huang, and Steven D. Blostein. "Least-squares fitting of two 3-D point sets." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 5 (1987): 698-700.