

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH**



CÔNG TRÌNH NGHIÊN CỨU KHOA HỌC CẤP TRƯỜNG

**ĐỊNH VỊ ROBOT DI ĐỘNG NGOÀI TRỜI BẰNG CAMERA
ĐA HƯỚNG KẾT HỢP
GLOBAL POSITIONING SYSTEM (GPS)**

MÃ SỐ: 2015-39TĐ



Tp. Hồ Chí Minh, tháng 03/2016

MỤC LỤC

MỤC LỤC	I
DANH SÁCH CÁC HÌNH	III
DANH SÁCH CÁC CHỮ VIẾT TẮT	V
Chương 1: TỔNG QUAN	1
1.1. Tổng quan chung về lĩnh vực nghiên cứu, các kết quả nghiên cứu trong và ngoài nước đã công bố.....	1
1.1.1. Tổng quan chung về định vị robot	1
1.1.2. Kết quả nghiên cứu trong và ngoài nước	4
1.2. Mục tiêu và đối tượng nghiên cứu.....	6
1.2.1. Mục tiêu nghiên cứu.....	6
1.2.3. Đối tượng nghiên cứu.....	6
1.3. Nhiệm vụ của đề tài và phạm vi nghiên cứu	6
1.3.1. Nhiệm vụ của đề tài.....	6
1.3.2. Phạm vi nghiên cứu	7
1.4. Phương pháp nghiên cứu	7
 Chương 2: Những công trình liên quan và đề xuất phương pháp định vị dùng omni-directional camera kết hợp GPS	
.....	Error! Bookmark not defined.
2.1. Tìm cặp điểm đặc trưng tương đồng	9

2.1.1. Các loại đặc trưng.....	9
2.1.2. Đặc trưng SIFT.....	12
2.2. Thông số nội camera.....	Error! Bookmark not defined.
2.3. Thuật toán RANSAC.....	17
2.4. Định vị camera dựa trên điểm tương đồng.....	Error! Bookmark not defined.
2.4.1. Hình học Epipolar và ma trận cơ bản.....	Error! Bookmark not defined.
2.4.2. Ma trận thiết yếu và thông số ngoại camera.....	Error! Bookmark not defined.
2.5. Giải thuật 8 điểm kết hợp RANSAC	Error! Bookmark not defined.
2.6. Thuật toán tối ưu xây dựng không gian đám mây điểm và vị trí robot.....	Error! Bookmark not defined.
Chương 3: THỰC NGHIỆM GIẢI THUẬT	32
3.1. Cân chỉnh camera tìm bộ thông số nội camera.....	Error! Bookmark not defined.
3.2. Trích rút đặc trưng SIFT	Error! Bookmark not defined.
3.3. Mô tả cặp đặc trưng tương đồng	Error! Bookmark not defined.
3.4. Ước lượng ma trận cơ bản F và ma trận thiết yếu E	Error! Bookmark not defined.
Chương 4: KẾT LUẬN	36
PHỤ LỤC A	37
TÀI LIỆU THAM KHẢO	38

DANH SÁCH CÁC HÌNH

Hình 1.1: Dừng xe ô tô được trang bị thiết bị định vị GPS để lập bản đồ	Error!
Bookmark not defined.	
Hình 2.1: Phương pháp tính DoG.	13
Hình 2.2: Tìm điểm đặc trưng từ các giá trị DoG.....	13
Hình 2.3. Mô tả điểm đặc trưng	16
Hình 2.4: Mô hình Pinhole camera	Error! Bookmark not defined.
Hình 2.5: Hệ tọa độ tương đương mô hình Pinhole	Error! Bookmark not defined.
defined.	
Hình 2.6: Quan hệ giữa tọa độ camera và tọa độ không gian	Error! Bookmark not defined.
not defined.	
Hình 2.7: Quan hệ giữa tọa độ camera và tọa độ không gian	Error! Bookmark not defined.
not defined.	
Hình 2.8: Tập dữ liệu đường thẳng	18
Hình 2.9: Tìm đường thẳng dùng thuật toán RANSAC.....	18
Hình 2.10: Mô tả hình học Epipolar	Error! Bookmark not defined.
Hình 2.11: Mô tả hình học của 3 frame liên tục với những đặc trưng tương đồng	Error! Bookmark not defined.
defined.	
Hình 3.1: Lưu đồ thực hiện tìm vị trí.....	Error! Bookmark not defined.
Hình 3.2: Tập ảnh bàn cờ dùng cân chỉnh camera	Error! Bookmark not defined.
Hình 3.3: Ảnh qua bộ lọc Gaussian	Error! Bookmark not defined.
Hình 3.4: Ảnh thể hiện sau khi tính DoG.	Error! Bookmark not defined.
Hình 3.5: Điểm đặc trưng của ảnh	Error! Bookmark not defined.

Hình 3.6: Hình thể hiện các đặc trưng SIFT **Error! Bookmark not defined.**

Hình 3.7: Cặp điểm đặc trưng tương đồng..... **Error! Bookmark not defined.**

Hình 3.8. Trích điểm đặc trưng SIFT và đối sánh **Error! Bookmark not defined.**

Hình 3.9. Tọa độ vị trí camera trong hai khung ảnh đầu tiên**Error! Bookmark not defined.**

Hình 3.10. Kết quả mô phỏng vị trí robot dùng 372 ảnh.**Error! Bookmark not defined.**

Hình 3.11. Kết quả thực nghiệm đám mây điểm 3D và vị trí robot**Error! Bookmark not defined.**

DANH SÁCH CÁC CHỮ VIẾT TẮT

SFM: Structure from motion.

DOF: degrees of freedom

IMU: inner measurement unit

SLAM: Simutaneous Localization and Mapping

BA: Bundle Adjustment

RANSAC: Random sample consensus.

SIFT: Scale invariant feature transform.

IR: Infrared radiation

GPS: Global Positioning System

Chương 1

TỔNG QUAN

1.1. Tổng quan chung về lĩnh vực nghiên cứu, các kết quả nghiên cứu trong và ngoài nước đã công bố

1.1.1. Tổng quan chung về định vị robot

Đề tài này đề xuất một phương pháp định vị cho các thiết bị di chuyển ngoài trời dùng thuật toán một điểm kết hợp với một vùng tương đồng đường biên của hệ thống thị giác đơn trong môi trường đô thị rộng lớn. Việc định vị 5 bậc tự do (five degrees of freedom-5 DOF) ước lượng bởi camera đơn-đa hướng sẽ cải thiện được những giả định di chuyển trong mặt phẳng của hầu hết các nghiên cứu trước đây. Trong những năm gần đây mô hình di chuyển tựa như xe hơi với giả định chuyển động trong mặt phẳng thì thường được nghiên cứu để giảm bớt những yêu cầu về điểm tương đồng cho đến còn một điểm. Tuy nhiên, trong những ứng dụng thực tế của việc di chuyển xa ở ngoài trời, những di chuyển này hầu như không thỏa mãn điều kiện này. Điều này dẫn đến sự không chính xác trong việc định vị. Trong phương pháp đề xuất này, mô hình tựa như xe hơi được tận dụng cho việc định vị 5 bậc tự do, tuy nhiên những yêu cầu về những điểm đặc trưng tương đồng thì được giảm xuống chỉ còn một điểm kết hợp với một vùng đối sánh, phương pháp này thì đơn giản hơn rất nhiều so với phương pháp 5 –điểm RANSAC. Những sai số tích lũy của hệ thống đo lường bằng hình ảnh sẽ được loại bỏ bằng cách kết hợp với hệ thống định vị GPS dưới dạng ước lượng cực đại hàm xác suất trong phương pháp lọc Kalman mở rộng. Những ứng dụng thực tế trên vùng đồi núi sẽ cho thấy độ chính xác của việc định vị theo phương pháp này.

Việc điều hướng cho các thiết bị tự hành hay rô bốt là một hướng nghiên cứu quan trọng có thể áp dụng trong nhiều lĩnh vực khác nhau như định vị, hoạch

định đường đi và đối sánh với bản đồ. Mặc dù một vài thành quả đã đạt được trong lĩnh vực này nhưng vẫn chưa có phương pháp nào thỏa mãn yêu cầu về độ chính xác cao cũng như là tính bền vững khi khoảng cách di chuyển xa trong các điều kiện địa hình hay môi trường khác nhau.

Trong những năm gần đây, nhiều phương pháp đã được phát triển cho việc điều hướng các thiết bị di chuyển có thể tạm chia ra một số phương pháp như sau: Phương pháp sử dụng chỉ những thiết bị điện tử (ví dụ như hệ thống định vị toàn cầu GPS, đơn vị đo lường bên trong (IMU), dụng cụ đo vòng quay của bánh xe, cảm biến laze), phương pháp chỉ dựa trên quan sát (camera đơn, stereo camera, camera đa hướng phản xạ qua gương) và phương pháp kết hợp giữa các thiết bị điện tử và hệ thống quan sát bằng hình ảnh. Trong nhóm đầu tiên, thiết bị GPS gắn trên thiết bị di chuyển nhận tín hiệu từ các vệ tinh và sau đó cho ra tín hiệu tuyệt đối về vị trí trên bản đồ. Độ chính xác của phương pháp này thường là thấp, tín hiệu từ vệ tinh thường bị dịch đi so với giá trị thực, hơn nữa các tòa nhà trong vùng đô thị thường gây nên hiện tượng làm mất hay nhảy tín hiệu từ vệ tinh vì vậy thiết bị di chuyển có thể bị mất thông tin. Để cải thiện phương pháp này một số cảm biến khác được sử dụng bổ sung, hệ thống sẽ trở thành hệ thống đa cảm biến. Ví dụ như việc bổ sung thêm thiết bị đo lường vòng bánh xe có thể cải thiện độ chính xác trong dịch chuyển nhưng nó có thể gây ra sai số nếu bánh xe bị trượt hay di chuyển trên những con đường gồ ghề. Loại thiết bị khác cũng được xem xét đến đó là IMU. Thiết bị này được dùng cho việc đo lường gia tốc và hướng. Tuy nhiên, giá thành của các loại thiết bị này khá là đắt tiền. Cảm biến laser cũng là một lựa chọn tốt trong những trường hợp này, tuy nhiên tín hiệu có thể bị yếu khi đối tượng xuất hiện ở khoảng cách xa hoặc không có phản xạ. Nhìn chung, tín hiệu GPS thì chính xác về hình dạng trong việc di chuyển rộng, toàn cục nhưng sẽ không chính xác trong các vị trí cục bộ. Đặc tính này rất quan trọng có thể được tận dụng cho việc định vị.

Trong nhóm thứ hai, hệ thống thị giác được gắn trên các thiết bị di chuyển và thu thập hình ảnh/ Video từ các cảnh vật xung quanh cho các bước xử lý sau

đó. Thông thường những ràng buộc về góc quay và tịnh tiến của những khung hình liên tục thì được phân tích trước tiên. Một số những đặc trưng bên vững được xem như là điểm mốc ví dụ như những điểm, đường thẳng, mặt phẳng. Sau khi hệ thống tìm ra sự tương đồng của những đặc trưng này, sự dịch chuyển của đối tượng được tính toán. Ở đây sự không rõ ràng về tính tỷ lệ rất quan trọng. Với những thiết bị hình ảnh đã được cân chỉnh, tỷ lệ của mô hình và cảnh vật thật sự sẽ được biết một cách rõ ràng. Tuy nhiên, tỷ lệ ban đầu phải được ước lượng dùng các thiết bị hỗ trợ ví dụ cảm biến laser kết hợp với camera đơn. Nhìn chung, những loại hệ thống như thế này có thể tạo ra kết quả chính xác trong khoảng di chuyển ngắn hay môi trường không có tín hiệu GPS. Vấn đề lớn ở đây là sai số tích lũy khi thiết bị di chuyển làm việc trong môi trường rộng lớn thì sai số tích lũy này sẽ lớn dần, Vì vậy, quy đạo cuối cũng sẽ bị phân kỳ khi so sánh với giá trị thực. Tuy nhiên một ưu điểm rõ ràng của phương pháp này là những sai số cục bộ trong khoảng di chuyển nhỏ là rất nhỏ. Đặc trưng này sẽ được tận dụng trong đề tài nghiên cứu này.

Trong một nhóm thứ ba, sự kết hợp của thiết bị điện tử và những phương pháp dựa trên thị giác được đề xuất. Những phương pháp loại này thì tận dụng ưu điểm của mỗi loại cảm biến. Tín hiệu toàn cục từ GPS được dùng làm thông tin tham khảo về vị trí của cảm biến thị giác. Hơn nữa, sự dịch chuyển của hệ thống thị giác được ước lượng một cách dễ dàng trong môi trường ngoài trời giàu thông tin (ví dụ như rất nhiều vật mốc là tự nhiên và nhân tạo). Hệ thống thị giác này có thể dùng thông tin toàn cục từ GPS để điều chỉnh lại vị trí hiện tại. Bên cạnh đó những cảm biến này có thể bù trừ lẫn nhau trong trường hợp mất một trong hai tín hiệu. Ví dụ như khi thiết bị di chuyển trong các đường hầm hay các thành phố có mật độ tòa nhà dày đặc, khi đó tín hiệu GPS có thể bị mất hoặc dao động, khi đó hệ thống thị giác có thể hoạt động độc lập. Mặc khác hệ thống thị giác có thể chỉnh định lại tính hiệu GPS bị sai. Công trình nghiên cứu này được tiến hành dựa trên những nhận định đó.

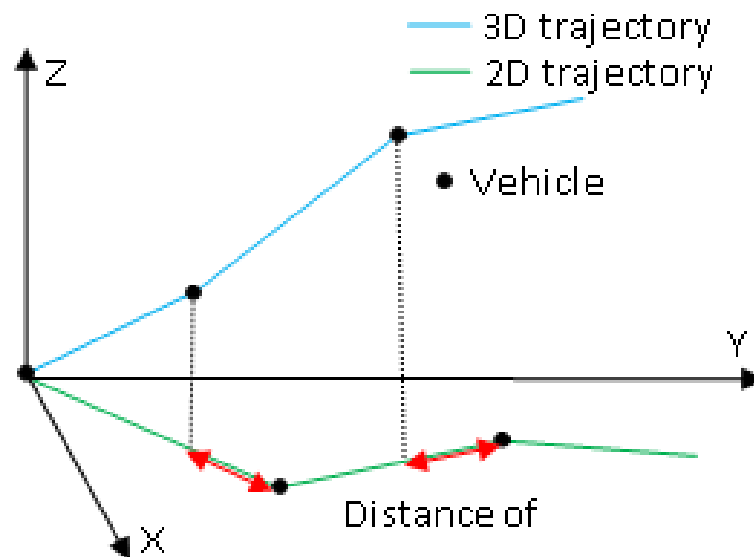
1.1.2. Kết quả nghiên cứu trong và ngoài nước

Lập bản đồ có nhiều cách thức khác nhau như dùng thiết bị định vị toàn cầu GPS gắn trên ô tô. Khi ô tô di chuyển, tọa độ hiện tại của ô tô luôn được cập nhật gửi về máy chủ. Từ đó máy chủ sẽ tính toán và vẽ được bản đồ 2D của môi trường mà ô tô di chuyển. Tuy nhiên, độ chính xác của thiết bị GPS không tốt, sai số từ vài mét cho đến vài chục mét.

Chế độ tự định vị là một trong những hướng nghiên cứu quan trọng của robot làm việc ngoài trời có thể ứng dụng trong việc tự hành, xây dựng bản đồ, hoạch định đường đi. Mặc dù nhiều nghiên cứu đã được tiến hành trong thời gian qua nhưng hầu như chưa có giải pháp nào tối ưu và bền vững khi robot làm việc lâu dài ngoài trời với khoảng cách xa và môi trường làm việc khác nhau.

Sự kết hợp giữa hệ thống thị giác và những thiết bị điện tử có thể xem như là giải pháp hiện hữu giải quyết vấn đề sai số tích lũy trong những năm gần đây. Về cơ bản những giải pháp này có thể tạm chia thành 3 nhóm như sau: Nhóm thứ nhất chỉ dùng hệ thống thị giác, nhóm thứ hai dùng GPS kết hợp với các cảm biến điện tử. Nhóm thứ ba kết hợp hệ thống thị giác và GPS. Trong nhóm thứ nhất, những nghiên cứu sớm về hệ thống tính toán đường đi robot dùng camera truyền thống được tiến hành đầu tiên bởi Nister [1] và Royer [2], trong khi một số nhóm nghiên cứu khác như [3], [4] dùng hệ thống camera đôi. Bởi vì sự hạn chế trong góc nhìn của camera, một số tác giả khác đề xuất dùng camera đa hướng (Omnidirectional camera). Hệ thống định vị dùng camera đa hướng tiêu biểu có thể kể đến trong [5-7]. Nguyên lý chính trong phương pháp này là tìm sự tương quan đặc trưng của chuỗi hình ảnh thu được, sau đó tính toán những ràng buộc này để tìm ra được sự dịch chuyển của robot. Vấn đề khó khăn gặp phải trong giải pháp này là khi số lượng đặc trưng tương đồng quá ít, khó tìm được những đặc trưng tương đồng chính xác khi loại bỏ những tương đồng sai. Chính điều này dẫn tới sai số lớn trong một số khu vực dịch chuyển đặc biệt của robot. Tuy nhiên trong những vùng làm việc với khoảng cách ngắn, quỹ đạo có thể đúng. Nhưng

khi Robot làm việc với khoảng di chuyển xa thì quỹ đạo hầu như sẽ phân kì nếu so sánh với đường đi thực tế. Đây cũng là thách thức lớn trong vấn đề định vị robot bằng phương pháp tang dần. Trong nhóm phương pháp thứ hai, Những cảm biến điện tử được kết hợp thành một hệ thống cảm biến để định vị cho robot như trong [8], [9]. Thông thường GPS sẽ được dùng để định vị quỹ đạo dịch chuyển chung toàn cục và cảm biến IMU hoặc thiết bị đo dịch chuyển bánh được dùng để tính toán di chuyển cục bộ. Phương pháp này cho quỹ đạo toàn cục gần đúng, hầu như đều bị trượt đi so với quỹ đạo thực. Đây cũng là thách thức trong phương pháp định vị mà không dùng thị giác. Trong những năm gần đây, phương pháp kết hợp hai giải pháp bên trên được xem như giải pháp tốt hơn để khắc phục những nhược điểm đã nêu. Một số nhóm tác giả tiêu biểu như [10-12] đề xuất phương pháp kết hợp thị giác và GPS. Kết quả cho thấy độ chính xác được cải thiện khá nhiều. Tuy nhiên các phương pháp này dựa trên tương quan đặc trưng trong hầu hết các bước xử lý. Điều này làm tăng thời gian xử lý đáng kể. Hơn nữa những nghiên cứu này đề giả định mặt di chuyển là phẳng hoàn toàn, điều này không đúng trong các ứng dụng thực tế. Khi Robot di chuyển trên đường dốc, quỹ đạo hoặc vị trí của robot sẽ không còn đúng nữa như trong hình 1.



Hình 1. Sai số vị trí của robot di chuyển trong đường dốc

Theo thống kê của tác giả hiện nay chưa có nhiều công trình trong nước nghiên cứu về định vị robot di động ngoài trời. Một số nhóm tác giả trong nước nghiên cứu vấn đề định vị cho robot di động dựa trên camera nhưng hầu hết các di chuyển này là ngắn hoặc trong nhà. Những đề tài đã thực hiện có thể liệt kê [13-20].

1.2. Mục tiêu, đối tượng nghiên cứu

1.2.1. Mục tiêu nghiên cứu

Mục tiêu đề tài này là định vị robot dựa vào thuật toán tối ưu các đặc trưng trên ảnh omni. Robot sẽ dùng camera chụp các ảnh 2D trong quá trình di chuyển để ước lượng vị trí camera được gắn trên robot, sau đó tính toán vị trí và hướng của robot trong không gian robot đang tương tác. Đồng thời kết hợp với thông tin từ GPS để chỉnh sửa lại nhưng vị trí sai số do bị nhiễu tác động.

1.2.2. Đối tượng nghiên cứu

- Robot hoặc xe chuyên dụng
- Hệ thống thị giác máy, GPS

1.3. Nhiệm vụ của đề tài và phạm vi nghiên cứu

1.3.1. Nhiệm vụ của đề tài

Để đạt được mục tiêu nghiên cứu cần thực hiện các nhiệm vụ sau:

- Tìm hiểu lý thuyết về lập bản đồ và định vị robot
- Tìm hiểu về xử lý ảnh: các phép toán trong xử lý ảnh, trích đặt trung từ ảnh, loại bỏ những đặt trung yếu hoặc sai
 - Trích xuất đặc trưng về đường biên trong ảnh
 - Thuật toán trích đặt trung SIFT từ một hình ảnh
 - Thuật toán RANSAC loại bỏ các đặt trung yếu của một hình ảnh

- Thuật toán tối ưu tái cấu trúc 3D xây dựng bản đồ 3D, xác định vị trí và góc quay camera để tìm vị trí và hướng của robot có kết hợp hệ thống GPS.

1.3.2. Phạm vi nghiên cứu

Đề tài nghiên cứu trên thiết bị di chuyển có người lái, tập ảnh 2D được tác giả thu thập trong quá trình di chuyển trên đường đi. Trên tập dữ liệu ảnh 2D đề tài mô phỏng tính toán xây dựng bản đồ 3D và định vị vị trí camera trong bản đồ này.

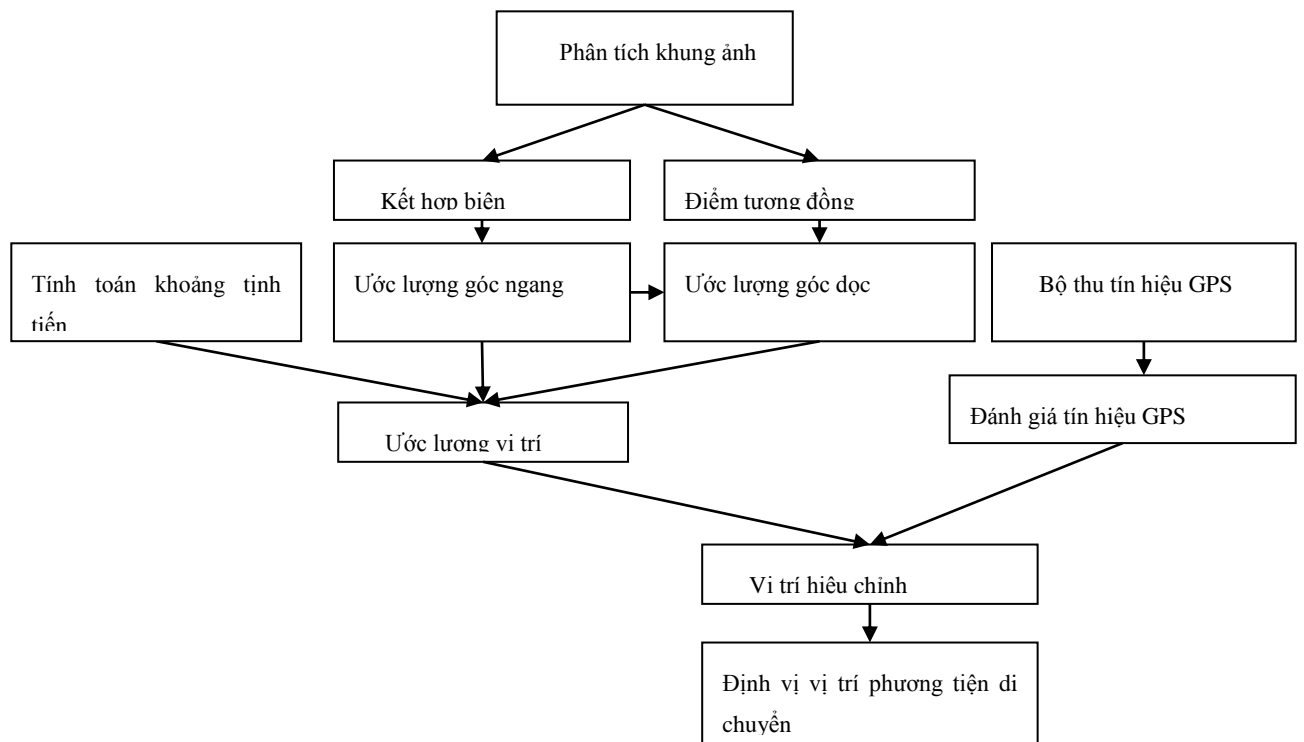
1.4. Phương pháp nghiên cứu

Cách tiếp cận của nghiên cứu này là khảo sát cơ sở lý thuyết những phương pháp hiện có, phân tích ưu nhược điểm. Tiến hành đề xuất phương pháp mới cải tiến. Thực hiện mô phỏng giải thuật trên phần mềm. Khi thu được sai số cho phép sẽ tiến hành thực nghiệm trên phần cứng và môi trường thực tế. Tóm lại, phương pháp nghiên cứu của đề tài là trước tiên tiến hành nghiên cứu lý thuyết, sau đó kiểm chứng bằng thực tế và rút ra kết luận.

Chương 2

Những công trình liên quan và đề xuất phương pháp định vị dùng omni-directional camera kết hợp GPS

Dựa trên những phân tích trên, phương pháp kết hợp camera đa hướng và thiết bị GPS tiêu chuẩn được đề xuất. Nhiều thông tin của môi trường xung quanh sẽ thu nhận được bằng camera có góc quan sát 360^0 , thậm chí khi thiết bị thực hiện một góc quay lớn thì điểm mốc vẫn được theo dõi. Đây là một ưu điểm rất lớn trong việc sử dụng camera cầu đa hướng với việc sử dụng thông tin dẫn đường từ GPS, quỹ đạo di chuyển được duy trì một cách chính xác vừa cục bộ từ hệ thống thị giác và hình dạng toàn cục từ hệ thống GPS. Sơ đồ tổng quát của phương pháp đề xuất như trên được trình bày trong hình số 2. Trong khung hình đầu tiên, hướng di chuyển của đối tượng được xác định. Sau đó những ma trận ràng buộc giữa các khung hình liên tục sẽ được tính toán. Những đặc trưng về đường biên và những điểm then chốt được kết hợp cho việc tính toán góc quay, ở đây là góc ngang và góc đứng. Phương pháp đối sánh Chamfer được sử dụng để tính toán góc ngang. Việc ước lượng góc ngang này tương tự như một la bàn thị giác. Bởi vì công trình này chỉ dựa trên chuyển động 5 bậc tự do bao gồm 3 chuyển động tịnh tiến và 2 giá trị góc quay nên góc đứng cũng cần phải được tính toán. Tối thiểu 1 điểm đặc trưng tương đồng thì cần thiết cho việc tính toán này. Hơn nữa, khoảng cách tịnh tiến của thiết bị là những giá trị tuyệt đối nhận được từ việc di chuyển đều của đối tượng. Khoảng cách tịnh tiến bằng nhau tương ứng với mỗi khoảng khung hình được trích ra. Trong giai đoạn tiếp theo, thiết bị GPS sẽ cung cấp vị trí toàn cục của đối tượng, thông tin này được dùng để hiệu chỉnh lại độ chính xác, Tuy nhiên, vị trí xác định từ GPS có thể không chính xác từ những môi trường có mật độ xây dựng lớn. Như vậy bước xử lý, đánh giá cần phải được thực hiện.



Hình 2: sơ đồ đề xuất về định vị vị trí phương tiện di chuyển

2.1. Tìm cặp điểm đặc trưng tương đồng

2.1.1. Các loại đặc trưng

Để tìm vị trí robot từ nhiều ảnh, việc tìm đặc trưng của ảnh cho phù hợp là thành phần quan trọng và thách thức. Việc lựa chọn các đặc trưng sẽ giúp cho việc tăng độ chính xác vị trí, hướng của robot và tốc độ xử lý của robot. Những năm gần đây việc sử dụng các đặc trưng cục bộ đã trở thành một hướng mới trong định vị robot bằng kỹ thuật xử lý ảnh. Các điểm đặc trưng này là bất biến trong phép biến đổi hình học, phép biến đổi affine.

Harris corner

Harris corner là đặc trưng phát hiện góc, hoặc một thuật ngữ tổng quát hơn là phát hiện điểm quan tâm (interest point detection) là một hướng tiếp cận được

sử dụng trong các hệ thống thị giác máy tính để trích chọn các loại đặc trưng và suy luận ra các nội dung của một ảnh. Việc phát hiện góc được dùng thường xuyên trong phát hiện, theo dõi chuyển động, mô hình 3D và nhận dạng đối tượng.

Một góc được xác định bởi nơi giao nhau của hai cạnh. Một góc cũng có thể được xác định như một điểm có hai hướng khác nhau trong một vùng cục bộ của điểm đó. Một điểm quan tâm là một điểm trong một ảnh mà điểm này có vị trí được xác định tốt và có thể được phát hiện nhanh chóng. Điều này có nghĩa là một điểm quan tâm có thể là một góc nhưng cũng có thể là một điểm đơn có giá trị cường độ cực đại hoặc cực tiểu cục bộ, các điểm kết thúc của đường thẳng hoặc một điểm trên một đường cong mà ở đó độ cong là tối đa cục bộ. Trên thực tế, hầu hết các phương pháp phát hiện góc phát hiện các điểm hơn là các góc nói riêng.

Phương pháp phát hiện góc Harris [5] là một phương pháp phát hiện điểm quan tâm phổ biến vì nó bất biến đối với phép quay, thay đổi độ sáng và tạp nhiễu ảnh. Phương pháp này dựa trên hàm tương quan tự động cục bộ của một tín hiệu; ở đó hàm tương quan tự động cục bộ đo các thay đổi cục bộ của tín hiệu với các mảnh ảnh được dịch chuyển một lượng nhỏ theo các hướng khác nhau. Tuy nhiên, phương pháp này chưa giải quyết được vấn đề biến đổi co, giãn ảnh hay những phép biến đổi affine.

Harris Laplace

Harris Laplace là thuật toán tìm điểm đặc trưng cải tiến từ thuật toán tìm điểm đặc trưng Harris. Thuật toán được đưa ra nhằm giải quyết vấn đề biến đổi theo tỉ lệ mà thuật toán Harris corner chưa khắc phục được. Harris Laplace đưa ra giải pháp có thể lựa chọn tỉ lệ tự động trong quá trình trích điểm đặc trưng từ ảnh. Thuật toán thực hiện dựa vào hai bước chính. Bước 1, thuật toán dùng hàm Harris Laplace để phát hiện điểm tỉ lệ. Bước 2, lựa chọn các điểm mà hàm

Laplace of Gaussian đạt giá trị cực đại. Đặc trưng Harris Laplace tuy giải quyết được vấn đề co giãn của ảnh, nhưng sự biến đổi affine vẫn còn hạn chế.

SURF

SURF là một phương pháp phát hiện và mô tả hình ảnh tốt, do Herbert Bay đưa ra lần đầu vào năm 2006, có thể được sử dụng trong các nhiệm vụ tầm nhìn của máy tính như nhận dạng đối tượng, phục hồi 3D. Nó dựa trên mô tả SIFT. Phiên bản tiêu chuẩn của SURF nhanh hơn nhiều so với SIFT và nó chống lại sự biến đổi hình ảnh khác nhau tốt hơn SIFT. SURF dựa trên tổng xấp xỉ các đặc trưng Haar Wavelet 2D và sử dụng hiệu quả cho các ảnh tích hợp (integral image).

Phương pháp này dựa trên các ma trận Hessian, nhưng sử dụng một xấp xỉ cơ bản, cũng giống như DoG là một phát hiện cơ bản dựa trên Laplacian. Nó dựa trên hình ảnh tích hợp để giảm thời gian tính toán và được gọi là phát hiện 'Fast-Hessian'. Mặt khác, mô tả này mô tả một phân bố các đặc trưng Haar-Wavelet trong các lân cận của điểm quan tâm. Hơn nữa, chỉ sử dụng kích thước 64, để giảm thời gian tính toán các đặc trưng và mô tả bộ tương đồng, đồng thời tăng độ tin cậy. Nó không chỉ làm tăng tốc độ tương đồng, mà còn tăng độ tin cậy của bộ mô tả đặc trưng này.

SIFT

SIFT [6] được trình bày đầu tiên bởi David G. Lowe năm 1999, đến năm 2004 chính David G. Lowe hoàn thiện giải thuật SIFT [7]. Thí nghiệm về thuật toán SIFT của ông về một hình ảnh sau khi co giãn, xoay hay biến đổi trong không gian affine thì đặt trưng SIFT hầu như không thay đổi.

SIFT cho số lượng đặc trưng tốt hơn các giải thuật còn lại, nên tác giả lựa chọn làm cơ sở để tìm điểm đặc trưng trong nghiên cứu này.

2.1.2. Đặc trưng SIFT

Thuật toán SIFT được mô tả qua những bước chính sau: tìm cực trị của không gian tỉ lệ, lọc và trích xuất điểm đặc trưng, gán hướng cho điểm đặc trưng và mô tả điểm đặc trưng.

Tìm cực trị của không gian tỉ lệ

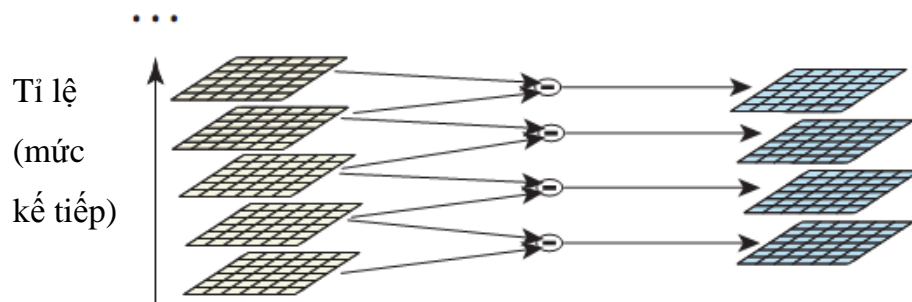
Đầu tiên, chúng ta xây dựng các mức của ảnh bằng phép biến đổi Gaussian của ảnh với các giá trị độ lệch chuẩn thay đổi liên tiếp nhau (gọi là ảnh Gaussian). Theo phương trình sau:

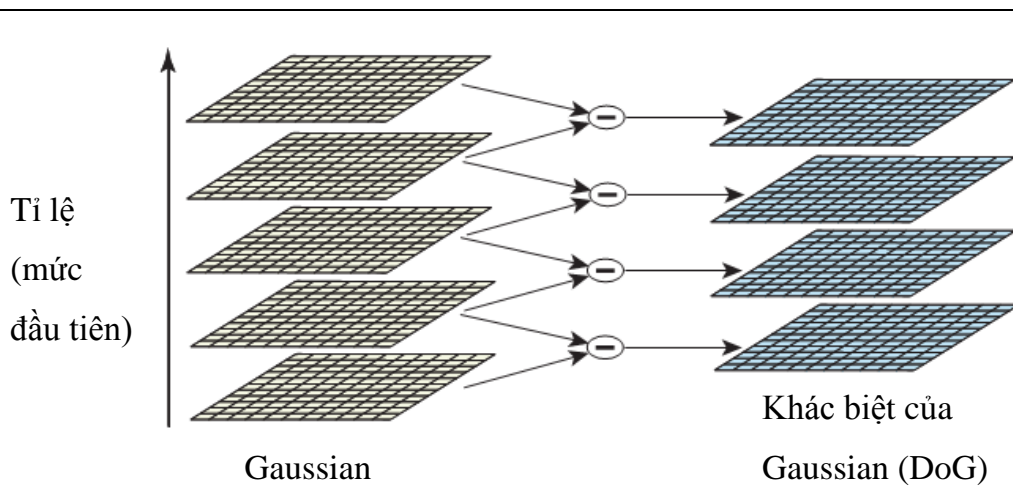
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad L(x, y, \sigma) = L(x, y, \sigma) * I(x, y) \quad (2.1)$$

Trong đó, dấu * là nhân chập trong miền x và y.

Sự khác biệt của Gaussian (DoG) trong mỗi cấp được tính bằng cách trừ hai ảnh Gaussian liền kề nhau.

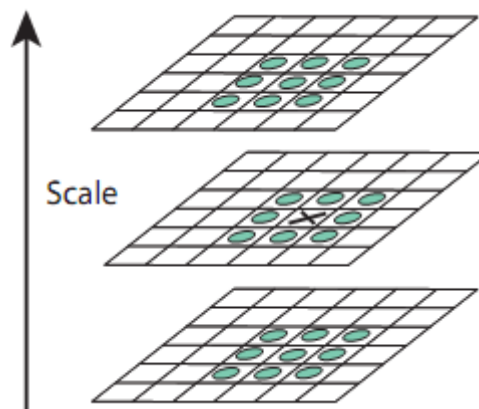
$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2.2)$$





Hình 2.1: Phương pháp tính DoG.

Trong Hình 2.1 ảnh bên trái thể hiện ảnh biến đổi Gaussian với các hệ số lệch chuẩn thay đổi, ảnh bên phải là kết quả trừ hai ảnh Gaussian kế nhau. Sau mỗi mức ảnh được giảm kích thước với tỉ lệ hệ số 2 và lặp lại quá trình [7].



Hình 2.2: Tìm điểm đặc trưng từ các giá trị DoG.

So sánh từng điểm ảnh trong tỉ lệ hiện tại với tỉ lệ trên, dưới trong lân cận 3x3, tức là 26 pixel xung quanh, chúng ta tìm giá trị lớn nhất hoặc nhỏ nhất giữa chúng. Điểm này được coi là điểm đặc trưng. Trong Hình 2.2 điểm đánh dấu x được so sánh với 26 điểm lân cận, nếu nó có giá trị lớn nhất hoặc nhỏ nhất thì được chọn làm điểm đặc trưng.

Lọc và trích xuất điểm đặc trưng

Sau khi tìm được điểm đặc trưng trong bước 2.1.1, không phải điểm đặc trưng nào cũng tốt nhất, vì vậy ta cần loại bỏ điểm đặc trưng yếu và trích xuất các điểm đặc trưng mạnh. Để thực hiện điều này ta cần thực hiện 3 bước sau: xác định đúng vị trí của điểm đặc trưng, loại bỏ điểm đặc trưng có tính tương phản thấp, loại bỏ điểm đặc trưng dư thừa theo biên.

- **Xác định đúng vị trí của điểm đặc trưng**

Vì điểm đặc trưng tìm được có thể chưa phải là điểm chính xác có giá trị cực đại hoặc cực tiểu. Ta dùng khai triển Taylor của hàm không gian tỉ lệ để dịch đến vị trí điểm lấy mẫu gốc.

$$D(X) = D + \frac{\partial D}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (2.3)$$

Trong đó: D và đạo hàm của nó được tính tại điểm đặc trưng và $X = (x, y, \sigma)^T$ là độ lệch của điểm này. Vị trí của điểm cực trị \hat{X} được tính bằng cách lấy đạo hàm của hàm trên theo X và cho nó tiến về 0.

$$\hat{X} = \frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (2.4)$$

- **Loại bỏ điểm đặc trưng có tính tương phản thấp**

Trong bước này ta sẽ loại bỏ những điểm đặc trưng có độ tương phản thấp. Để loại điểm đặc trưng có độ tương phản thấp ta so sánh giá trị $D(\hat{x})$ với giá trị ngưỡng. Với giá trị $D(\hat{x})$ được xác định bằng cách thế phương trình (2.5) vào phương trình (2.4), ta được:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^{-1}}{\partial X} \hat{X} \quad (2.5)$$

Nếu giá trị $D(\hat{x})$ thấp hơn ngưỡng thì điểm này bị loại trừ.

- **Loại bỏ điểm đặc trưng dư thừa theo biên**

Sau khi dùng bộ lọc Gaussian sẽ làm cho đường biên không còn rõ, các giá trị DoG ở dọc góc đường biên không thể hiện chính xác điểm đặc trưng, vì vậy chúng ta cần loại bỏ bớt các điểm đặc biệt trên biên. Chúng ta sử dụng ma trận Hessian 2x2 tính ở vị trí và hệ số tỉ lệ của điểm đặc trưng để tìm độ góc. Với công thức tỉ lệ của nguyên tắc góc:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2.6)$$

$$\frac{(D_{xx} + D_{yy})^2}{D_{xx}D_{yy} - (D_{xy})^2} < \frac{(r+1)^2}{r} \quad (2.7)$$

Nếu bất đẳng thức (2.8) không thỏa thì điểm đặc trưng bị loại bỏ [8].

Gán hướng cho điểm đặc trưng

Mỗi điểm đặc trưng được mô tả bởi hướng chính dựa vào vị trí của ảnh để biểu diễn đặc trưng này không biến đổi trong trạng thái xoay. Bước này được mô tả bởi hai phương trình sau:

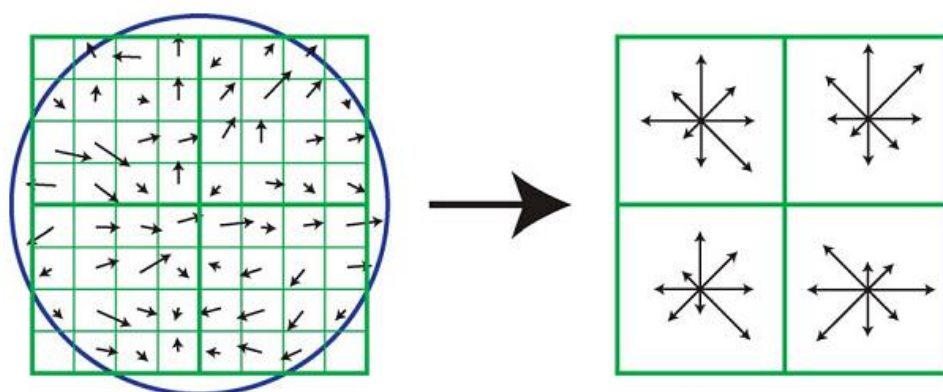
$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2.8)$$

$$(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (2.9)$$

Hai phương trình trên thể hiện độ lớn của Gradient và hướng của điểm pixel(x,y) tại L(x,y). Trong tính toán thực tế người ta xây dựng một Histogram Gradient xung quanh điểm đặc trưng từ các giá trị hướng của Gradient. Lược đồ này gồm 36 mức, thể hiện cho phạm vi 360^0 , vì vậy mỗi hướng có 10o, ta có tất cả 36 hướng. Mỗi mẫu thì được thêm vào trọng số của Histogram được tính từ độ lớn Gradient và tỉ lệ của điểm đặc trưng. Nếu Histogram nào có giá trị lớn hơn 80% của những giá trị khác thì được chọn làm hướng chính của điểm đặc trưng. Một điểm đặc trưng có thể có nhiều hướng chính [8].

Mô tả điểm đặc trưng

Tạo lại bảng mô tả 2D cho những giá trị trước đó thể hiện 3 thông số vị trí, tỉ lệ và hướng được dùng để mô tả một vùng của ảnh mà các giá trị này không thay đổi. Trong bước này tính toán bộ mô tả điểm đặc trưng. Mỗi điểm đặc trưng được mô tả bởi một vùng ô vuông 16×16 . Sau đó được chia làm 4 vùng nhỏ có kích thước 4×4 , tính giá trị Histogram với 8 mức hướng chính. Sau đó thêm vào giá trị độ lớn của gradient của vùng 4×4 vào cho các Histogram của hướng. Mỗi điểm đặc trưng chứa 8 vector, do đó một bộ mô tả sẽ chứa tổng cộng $4 \times 4 \times 8$ yếu tố.



Hình 2.3. Mô tả điểm đặc trưng. (a) Gradient của ảnh,
(b) mô tả điểm đặc trưng

Đôi sánh điểm đặc trưng

Những điểm đặc trưng tương đồng này được tìm ra bằng cách so sánh khoảng cách Eclidian của những vector đặc trưng. Điểm đặc trưng tương đồng là điểm có khoảng cách nhỏ nhất. Tuy nhiên khi số lượng điểm đặc trưng rất lớn hoặc phải so sánh với nhiều dữ liệu hình ảnh số lượng điểm đối sánh và phép so sánh sẽ rất lớn. Để khắc phục điều này, phương pháp cải tiến giải thuật cây so sánh K-D được áp dụng để giảm thời gian tính toán và so sánh. Trong phương pháp đối sánh này vẫn còn những lỗi do nhiễu của hình ảnh, phương pháp khác phục dựa trên thuật toán RANSAC được trình bày ở phần sau.

2.2. Thuật toán RANSAC

RANSAC (RANdom SAmple Consensus) là thuật toán được đưa ra bởi Fischler và Bolles [9], là một phương pháp ước lượng tham số mô hình, là phương pháp được thiết kế nhằm giải quyết một tỉ lệ lớn các giá trị ngoại lai.

Với những kỹ thuật thông thường người ta sử dụng càng những nhiều dữ liệu càng tốt để tạo thông số mô hình sau đó mới loại bỏ dần các điểm ngoại lai. Ở đây RANSAC sử dụng tập mẫu nhỏ nhất có thể để tìm thông số mô hình mà các dữ liệu điều thỏa mô hình trong khoảng cho phép.

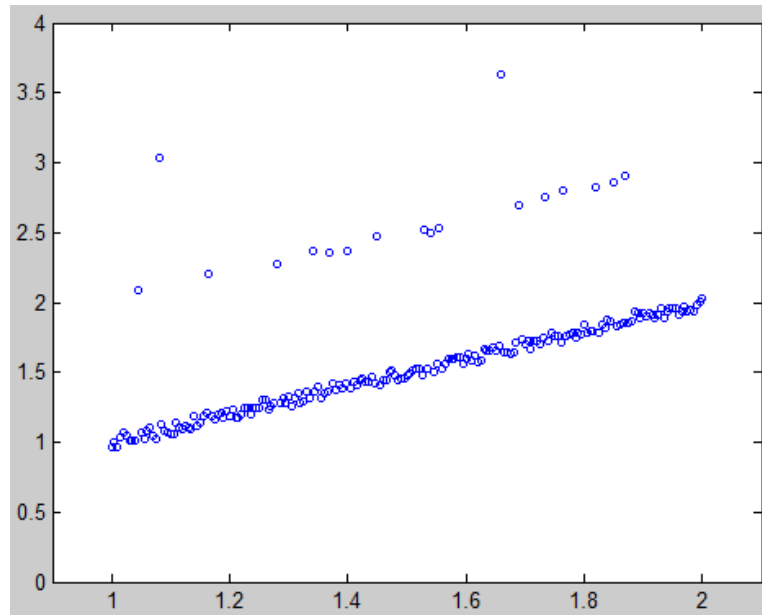
Thuật toán RANSAC:

1. Chọn ngẫu nhiên số lượng tối thiểu để xác định mô hình cần thiết
2. Giải tìm các thông số của mô hình
3. Xác định có bao nhiêu điểm trong tập dữ liệu thỏa mô hình trong khoảng cho phép (inlier)
4. Nếu tỉ số inlier trên tập dữ liệu nhỏ hơn ngưỡng τ thì lặp lại chọn dữ liệu khác tìm thông số của mô hình
5. Nếu không, lặp lại bước 1 đến 4 tối đa N lần

Vận dụng tìm đường thẳng $ax + by + c = 0$ bằng thuật toán RANSAC: Cho tập dữ liệu như Hình 2.8, tìm đường thẳng $ax + by + c$ để có 90% số điểm có khoảng cách đến đường thẳng nhỏ hơn một khoảng $k = 0.1$

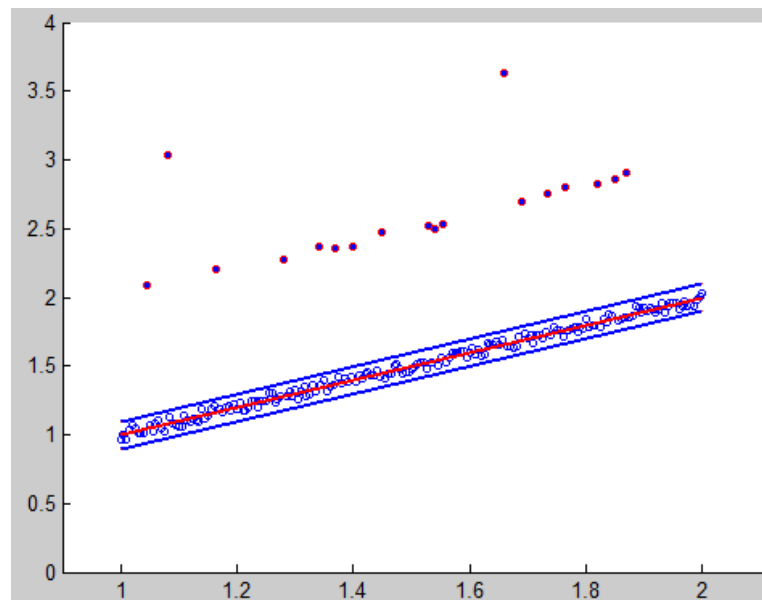
1. Chọn ngẫu nhiên 2 điểm (x_1, y_1) và (x_2, y_2)
2. Giải tìm các thông số a, b, c
3. Tính khoảng cách các điểm đến đường thẳng nếu nhỏ hơn 0.1 thì gọi là inlier
4. Tính tỉ số inlier trên tập dữ liệu nhỏ hơn nếu nhỏ hơn 90% thì lặp lại chọn (x_1, y_1) và (x_2, y_2) khác tìm lại thông số a, b, c

5. Nếu không, lặp lại bước 1 đến 4 tối đa N lần



Hình 2.8: Tập dữ liệu đường thẳng

Mô hình đường thẳng được tìm bằng thuật toán RANSAC thể hiện ở Hình 2.9 như sau:



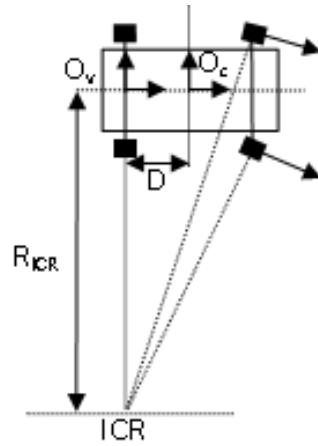
Hình 2.9: Tìm đường thẳng dùng thuật toán RANSAC

2.3. Hệ Thống Đo Lường Bằng Thị Giác Dựa Trên Camera Đa Hướng

Mục đích chính của giai đoạn này là tìm ra quỹ đạo chỉ bằng việc dùng hệ thống thị giác. Để làm việc đó, những ràng buộc của những khung hình liên tục phải được tính toán. Có 2 thông số quan trọng cần phải nhận được là khoảng tịnh tiến và góc quay. Đây cũng là một đề tài kinh điển trong việc nghiên cứu hình học đa góc của những thập niên trước. Một số tác giả trước đây đề xuất phương pháp dựa trên những đặc trưng để tính toán những ràng buộc, trong khi đó một số khác đề xuất phương pháp những đặc trưng về vùng đối tượng. Khoảng cách dịch chuyển tịnh tiến tuyệt đối thì được tính toán một cách tiêu biểu bằng việc sử dụng các thiết bị điện tử ví dụ như bộ mã khóa, vòng quay. Ngược lại khoảng cách dịch chuyển tịnh tiến trong phương pháp này được tính toán trực tiếp từ khoảng khung hình vì tốc độ di chuyển được giữ cố định.

2.3.1. Phân tích mô hình di chuyển

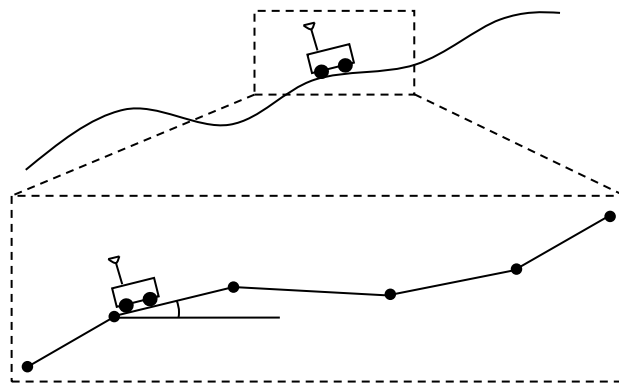
Trong nghiên cứu này, mỗi ràng buộc về dịch chuyển tuân thủ theo một hướng tròn gây ra bởi cấu trúc tương tự như xe hơi. Mô hình động học này được mô tả bởi Ackermann. Tâm của vòng trong được xem như là tâm của vòng quay tạm thời. Vì vậy, những con số rời rạc của những vòng quay xung quanh những ICR khác nhau tạo nên quỹ đạo di chuyển. Vị trí của camera trên máy của thiết bị cần phải được xem xét. Nếu như vị trí của camera trùng với trục của bánh sau thì nó sẽ có cùng hệ tọa độ với đối tượng. Nếu khoảng cách từ vị trí camera đến trục bánh sau khác thì một phép biến đổi hệ tọa độ camera sang hệ tọa độ của đối tượng cần phải được thực hiện. Hình ảnh minh họa được trình bày trong hình 3.



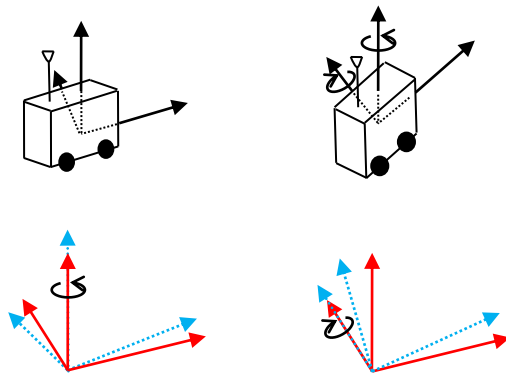
Hình 3: Mô hình Ackermann

Sự dịch chuyển của đối tượng trong môi trường thực tế thì không luôn luôn phẳng. Như trong hình 4 một vùng di chuyển khi được phóng lớn thì được mô tả bởi những di chuyển phẳng liên tục nhưng với một góc nghiêng khác nhau, vì thế trong mỗi di chuyển phẳng, vị trí của đối tượng được phân tích trong 5 DOF, ở đây bỏ qua góc cuộn. Hình số 5 minh họa các phân tích trên.

Ở đây, sự thể hiện góc quay và phép tịnh tiến dùng mô hình dịch chuyển của nhiều mặt phẳng được suy ra. Ký hiệu R và T là ma trận quay và vector tịnh tiến chưa biết của 2 khung hình liên tục.



Hình 4: Phân tích độ dốc của đường đi



Hình 5: Phân tích góc xoay: (a) trục gắn liền với phương tiện chuyển động, (b) góc dọc và ngang của phương tiện di chuyển, (c) góc xoay dọc, (d) góc xoay ngang

$$R = \begin{bmatrix} \cos \theta \cos \psi & \cos \varphi \sin \psi + \sin \varphi \sin \theta \cos \psi & \sin \varphi \sin \psi - \cos \varphi \sin \theta \cos \psi \\ -\cos \theta \sin \psi & \cos \varphi \cos \psi - \sin \varphi \sin \theta \sin \psi & \sin \varphi \cos \psi + \cos \varphi \sin \theta \sin \psi \\ \sin \theta & -\sin \varphi \cos \theta & \cos \varphi \cos \theta \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \quad (2.10)$$

ở đây $R = R_z R_y R_x$

Ghi chú 1: Cần phải chú ý khi thành lập ma trận quay. Phép nhân ma trận quay tương ứng theo trục x y z thì không có tính giao hoán trong đại số. Người thành lập có thể áp đặt thứ tự này ví dụ như xyz, xzy, và yzx là khác nhau. Ở đây, hệ thống theo tay phải được lựa chọn bởi vì nó thuận tiện để sử dụng trong hầu hết các ngành kỹ thuật và vật lý

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \quad (2.11)$$

$$R_Y = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix} \quad (2.12)$$

$$R_Z = \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.13)$$

Các góc quay có thể được suy ra từ ma trận quay:

$$\varphi = -\arctan(R_{32} / R_{33}) \quad (2.14)$$

$$\theta = \arcsin(R_{31}) \quad (2.15)$$

$$\psi = -\arctan(R_{21} / R_{11}) \quad (2.16)$$

Vector dịch chuyển tịnh tiến trên hệ trục tọa độ 3 chiều được tính toán dựa trên hình số 6 như sau:

$$T = \rho \begin{bmatrix} \cos \gamma \cos \alpha \\ \sin \gamma \\ \cos \gamma \sin \alpha \end{bmatrix} = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \quad (2.17)$$

Ngược lại với phương pháp truyền thống sử dụng 5 điểm tương đồng cho ước lượng vị trí của camera 6 bậc. Phương pháp này cần chỉ một điểm tương đồng công với một vùng đối xứng đường biên bằng cách dùng mô hình dịch chuyển. Những ràng buộc đặc biệt về mặt hình học sẽ giảm số biến trong ma trận quay và vector tịnh tiến. Ở đây giả định rằng độ nghiêng của đường chỉ theo một phương đúng, vì vậy góc cuộn thay đổi rất nhỏ và nó được bỏ qua. Ma trận quay được viết lại dưới hình thức đơn giản hơn khi góc quay không thay đổi ($\phi=0$)

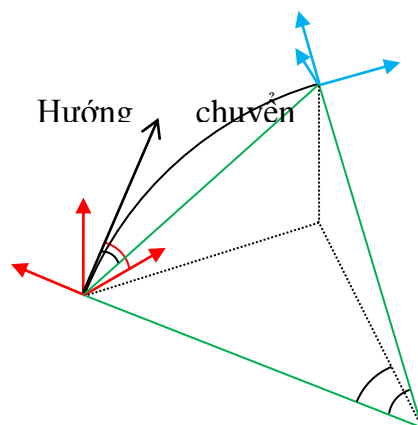
$$R_{YZ} = \begin{bmatrix} \cos \theta \cos \psi & \cos \theta \sin \psi & -\sin \theta \\ -\sin \psi & \cos \psi & 0 \\ \sin \theta \cos \psi & \sin \theta \sin \psi & \cos \theta \end{bmatrix} \quad (2.18)$$

Trong cấu trúc phần cứng , camera được cố định trên máy của đối tượng, vì vậy góc nghiêng α của mỗi mặt phẳng sẽ bằng góc θ . Hơn nữa, dựa theo những phân tích hình học trong phụ lục A hình 15, mối quan hệ giữa β và ψ thì nhận được như sau:

$$\frac{\beta}{\psi} = \frac{1}{\cos \alpha} \quad (2.19)$$

Ma trận quay được viết lại như sau:

$$R_{YZ} = \begin{bmatrix} \cos \alpha \cos(\beta \cos \alpha) & \cos \alpha \sin(\beta \cos \alpha) & -\sin \alpha \\ -\sin(\beta \cos \alpha) & \cos(\beta \cos \alpha) & 0 \\ \sin \alpha \cos(\beta \cos \alpha) & \sin \alpha \sin(\beta \cos \alpha) & \cos \alpha \end{bmatrix} \quad (2.20)$$



Hình 6: Mô hình chuyển động cung tròn

Ghi chú 2: Ma trận quay này chỉ phụ thuộc vào 2 biến số α và β , vector dịch chuyển cũng được tính toán.

Dựa trên mô hình chuyển động tròn khi $\gamma = \beta/2$, ma trận tịnh tiến được viết lại như sau

$$T = \rho \begin{bmatrix} \cos \gamma \cos \alpha \\ \sin \gamma \\ \cos \gamma \sin \alpha \end{bmatrix} = \rho \begin{bmatrix} \cos(\beta/2) \cos \alpha \\ \sin(\beta/2) \\ \cos(\beta/2) \sin \alpha \end{bmatrix} = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \quad (2.21)$$

Ghi chú 3: Vector tịnh tiến này cũng phụ thuộc vào 2 biến số α và β . Giá trị ρ trong nghiên cứu này được giữ làm hằng số khi tốc độ của xe được giả sử là không đổi.

2.3.2. Hệ thống đo lường bằng hình ảnh

Hệ thống đo lường bằng hình ảnh được tạo ra từ những ràng buộc về các cặp ảnh liên tục. Những loại ràng buộc này được phân tích trực tiếp từ ràng buộc epipolar thông qua ma trận cơ bản. Điều này được mô tả trong công thức sau:

$$p'^T E p = 0 \quad (2.22)$$

ở đây p và p' là những điểm tham chiếu hình ảnh trên hệ tọa độ cầu. Ma trận này được tạo ra dựa trên phép nhân của vector chuyển vị và ma trận quay, do đó

$$E = [T]_{\mathcal{K}} R = [T]_{\mathcal{K}} R_{YZ} \quad (2.23)$$

ở đó:

$$\begin{aligned} [T]_{\mathcal{K}} &= \begin{bmatrix} 0 & -T_Z & T_Y \\ T_Z & 0 & -T_X \\ -T_Y & T_X & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\cos(\beta/2)\sin\alpha & \sin(\beta/2) \\ \cos(\beta/2)\sin\alpha & 0 & -\cos(\beta/2)\cos\alpha \\ -\sin(\beta/2) & \cos(\beta/2)\cos\alpha & 0 \end{bmatrix} \end{aligned} \quad (2.24)$$

Một vài phương pháp thường được áp dụng để ước lượng hình ảnh của camera. Trong số đó, thì việc áp dụng thuật toán 5 điểm như mô tả trong [14] là một trong những phương pháp cổ điển. Gần đây, việc áp dụng 1 điểm Ransac [5] được xem như là phương pháp điển hình nhằm giảm số biến tương ứng dựa trên mô hình chuyển động tựa như xe hơi hoặc các phương tiện di chuyển khác. Tuy nhiên điểm giới hạn của phương pháp này là phải dựa trên giả thuyết các chuyển động là hai chiều. Giả thuyết này không hoàn toàn đúng so với thực tế. Như đã đề cập ở trên, khi một phương tiện di chuyển trên một địa hình không phẳng, thì việc thiết lập quỹ đạo chuyển động phẳng sẽ không còn chính xác nữa. Để khắc

phục vấn đề này, thì một phương pháp khác được đề xuất dựa trên mô hình chuyển động ba chiều. Từ ma trận cơ bản, phải có ít nhất 2 điểm tương ứng để giải phương trình (13). Tuy nhiên để giảm số vòng lặp RANSAC, phương pháp này đã sử dụng đặc tính bên ngoài để tính toán một biến. Nói cách khác, đặc tính bên ngoài được dùng để tính góc beta. Biến còn lại được tính toán dựa trên phương pháp 1 điểm RANSAC [5]

1) Tính toán góc β : Phương pháp đối sánh chamfer [15] được áp dụng để tính vùng tương đồng đường biên của các ảnh đa hướng. Điều này tương tự như hoạt động của một la bàn thị giác

Vùng được trích ra theo hướng phía trước của khung ảnh đầu tiên và thứ hai sẽ được xem như là vùng tương thích mẫu (ROI). Điều này được mô tả trong hình 7. Thứ nhất, hai mẫu vùng tương thích này sẽ được lấy ra từ thuật toán tách vùng biên. Ở đây sẽ dùng thuật toán Canny. Thứ hai, khoảng cách biên đôi của 2 vùng tương thích này sẽ được tính toán. Đầu tiên, độ phân giải trên vùng biên được gán với 0 và vùng còn lại được tính ở vô cùng. Phương pháp tính khoảng cách Euclidean được dùng để tính khoảng cách di chuyển này.

$$D_{i,j}^k = \min(D_{i-1,j}^{k-1} + 1, D_{i-1,j-1}^{k-1} + \sqrt{2}, D_{i,j-1}^{k-1} + 1, D_{i+1,j-1}^{k-1} + \sqrt{2}, D_{i+1,j}^{k-1} + 1, D_{i+1,j+1}^{k-1} + \sqrt{2}, D_{i,j+1}^{k-1} + 1, D_{i-1,j+1}^{k-1} + \sqrt{2}, D_{i,j}^{k-1}) \quad (2.25)$$

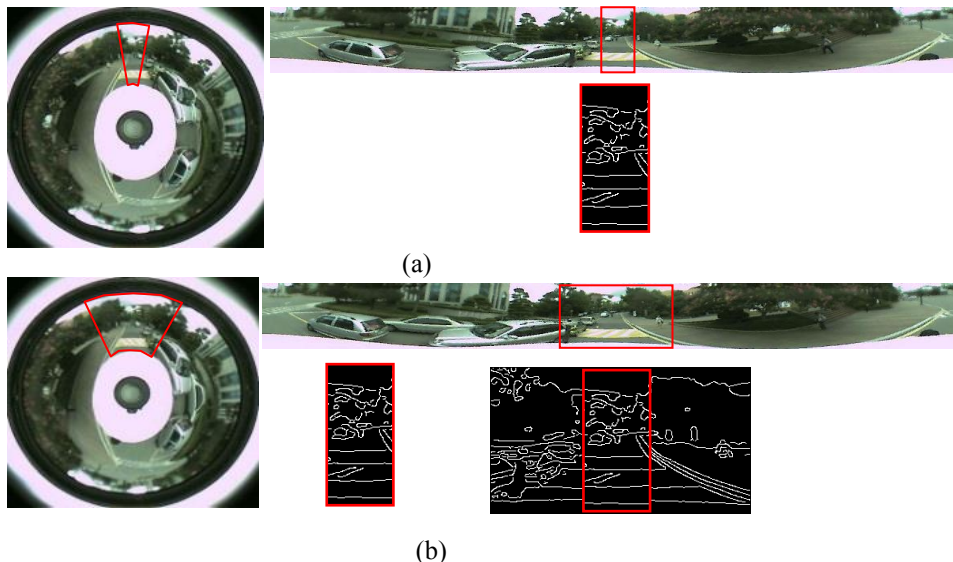
Trong đó $D_{i,j}^k$ là độ phân giải tại vị trí (i, j) của vòng lặp k. Vòng lặp này được thực hiện liên tục cho đến khi giá trị của độ phân giải không thay đổi. Vùng mẫu được tính toán bởi ảnh nhị phân với vùng biên được gán với giá trị 1 và 0. Thứ ba là tính toán tương tự giữa mẫu và vùng tương thích. Ở đây là tính toán tổng giá trị độ phân giải mà vùng biên trùng với vùng tương thích DTROI. Tính dọc theo các mẫu trên DTROI để tìm ra sự kết hợp tối ưu nhất với giá trị tổng nêu trên là nhỏ nhất. Sự kết hợp hoàn hảo nhất là giá trị này tiến đến 0, khi

mà vùng mẫu nằm chồng lên vùng ROI. Khi việc kết hợp này được tìm thấy thì hướng của khung hình tiếp theo sẽ được mô tả trong hình số 8.

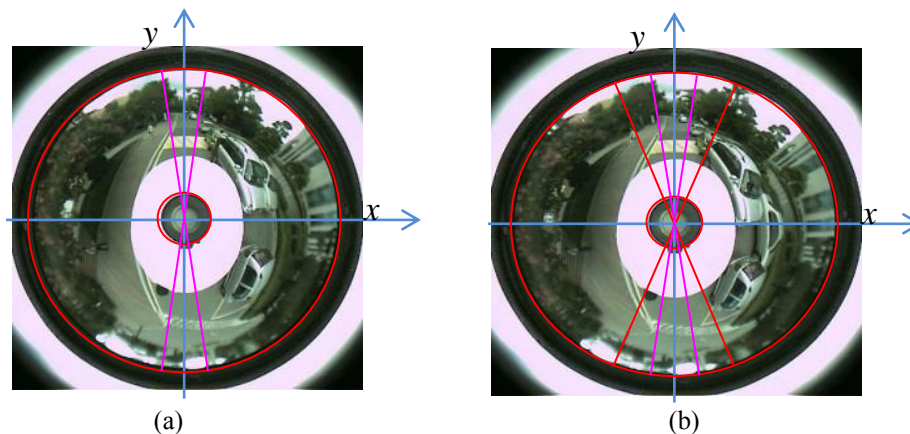
Bây giờ ma trận cơ bản chỉ gồm 1 biến. Phương pháp giải quyết biến này sẽ được giới thiệu trong phần dưới đây:

2) Tính toán góc α : Đặc tính SIFT [16] được trích ra từ những hình ảnh đa hướng mô tả trong hình số 9. Quá trình kết hợp này được mô tả trong thuật toán RANSAC.

Thuật toán này không phụ thuộc vào loại camera, các đặc tính trên hình ảnh được tham chiếu trên mặt phẳng cầu. Ý tưởng này sẽ thực hiện được nếu camera được hiệu chỉnh. Như đã đề cập ở [17], có 4 bước để thể hiện tham chiếu này và được thể hiện ở hình số 10



Hình 7: Kết hợp vùng biên (a) vùng hướng phía trước (b) vùng kết hợp ở khung hình tiếp theo



Hình 8: Hướng xoay: (a) hướng của khung hình phía trước, (b) hướng đã tính toán cho khung hình tiếp theo.



Hình 9: Sự kết hợp dựa trên đặc tính SIFT

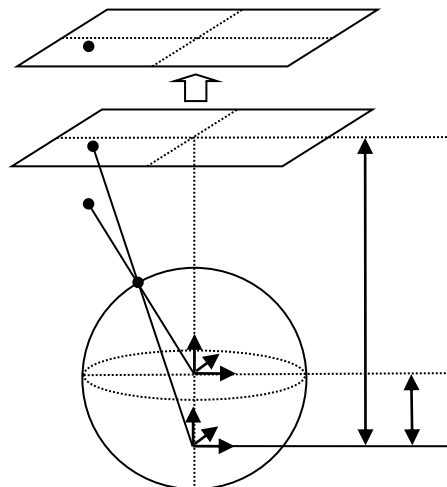
Bước đầu tiên là tham chiếu các điểm ảnh trên tọa độ cầu:

$$P_s = \frac{P}{\|P\|} = (x_s, y_s, z_s)^T \quad (2.26)$$

Bước thứ hai là chuyển trục của tọa độ cầu về một tọa độ tham chiếu mới với tâm là $C_\xi = (0,0,-\xi)$:

$$P_\xi = (x_s, y_s, z_s + \xi)^T \quad (2.27)$$

Khoảng cách ξ thay đổi từ 0 đến 1 và phụ thuộc vào loại kính. Khoảng cách sẽ bằng 0 khi một trong những bề mặt kính là phẳng hoặc parabol tương ứng. Với những loại kính khác, giá trị này được tính bằng khoảng cách d giữa tiêu điểm và đường chuẩn 1.



Hình 10: Mô hình camera đa hướng hình cầu

Giá trị ξ được mô tả trong bảng 1 như sau:

Loại kính	ξ	η
Parabol	1	$-2l$
Hyperbol	$\frac{d}{\sqrt{d^2 + 4l^2}}$	$\frac{-2l}{\sqrt{d^2 + 4l^2}}$
Elip	$\frac{d}{\sqrt{d^2 + 4l^2}}$	$\frac{2l}{\sqrt{d^2 + 4l^2}}$
Tương ứng	0	-1

Bước thứ ba là tham chiếu từ C_ξ lên mặt phẳng ảnh thông thường

$$m = (x_m, y_m, 1)^T = \left(\frac{x_s}{z_s + \xi}, \frac{y_s}{z_s + \xi}, 1 \right)^T = g^{-1}(P_s) \quad (2.28)$$

Bước thứ tư là tham chiếu từ mặt phẳng ảnh thông thường lên hình ảnh của camera thông qua ma trận K :

$$p = Km \quad (2.29)$$

Ở đây,

$$K = \begin{bmatrix} f & f\alpha & u_0 \\ 0 & fr & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.30)$$

Với f là tiêu cự, (u_0, v_0) điểm cơ bản, α là thông số skew và r là hệ số biến dạng

Cuối cùng ma trận biến đổi g từ điểm m đến điểm tương ứng trên mặt cầu được mô tả như sau:

$$g(m) = \begin{bmatrix} \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} x \\ \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} y \\ \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} - \xi \end{bmatrix} \quad (2.31)$$

Để tìm ra đối sánh đúng giữa các điểm tương ứng thì phương pháp 1-points RANSAC được áp dụng [5]

2.4. SỰ HIỆU CHỈNH SAI SỐ ĐO VỊ TRÍ CỦA HỆ THỐNG DẪN ĐƯỜNG QUA VỆ TINH GPS

Có nhiều phương pháp cho việc kết hợp nhiều cảm biến đo. Ở đây việc hiệu chỉnh từ thiết bị GPS được mô tả chỉ trên trục XY, Bởi vì sai số trên trục Z của hệ thống đo lường hình ảnh là nhỏ trong khi sai số do thiết bị GPS là lớn. Việc khắc phục các sai số này sẽ dựa trên phương trình dưới dạng ước lượng cực đại hàm xác suất trong phương pháp lọc Kalman mở rộng.

2.4.1. Đánh giá giá trị đo từ GPS

Trong tất cả các hệ thống định vị vị trí ở khoảng rộng, thì GPS được xem là một trong những thiết bị đo chính có định hướng. Tuy nhiên để có được độ chính xác của thiết bị này đòi hỏi phải rất tốn kém. Với một đầu thu chuẩn, thì tín hiệu thu từ GPS là không chắc chắn về độ tin cậy. Do đó việc đánh giá các dữ liệu ban đầu sẽ được thể hiện ở hình 11. Đặc biệt là khi một thiết bị di chuyển trên một vùng nhiều mật độ di chuyển hoặc trong khu đô thị thì tín hiệu sẽ bị phản xạ, phân tán hoặc thậm chí bị lạc hướng. Điều này dẫn đến sai số trong việc xác định vị trí dùng GPS. Vì lý do này mà các vị trí đo từ GPS với sai số nhỏ sẽ được dùng để hiệu chỉnh thiết bị đo bằng hình ảnh. Còn các sai số lớn sẽ bị loại bỏ. Hơn nữa việc ước lượng độ cao của phương tiện di chuyển trên

$$r = \begin{cases} \frac{\operatorname{tg}(2\varphi)(\sigma_x^2 - \sigma_y^2)}{2\sigma_x\sigma_y}, 0 < \varphi < \frac{\pi}{2} \\ \frac{\operatorname{tg}(2\varphi)(\sigma_y^2 - \sigma_x^2)}{2\sigma_x\sigma_y}, -\frac{\pi}{2} < \varphi < 0 \end{cases} \quad (2.34)$$

Nếu độ lệch chuẩn x và y là thấp hơn so với ngưỡng đưa ra trước đó thì thiết bị GPS này đạt độ chính xác cao dẫn đến thiết bị đo bằng hình ảnh cũng chính xác. Nếu không chỉ dùng kết quả từ việc đo bằng hình ảnh để xác định vị trí của phương tiện di chuyển.

2.4.2. Chuẩn hóa sai số đo từ GPS:

EKF được sử dụng rộng rãi trong việc xác định vị trí từ hệ thống đo lường kết hợp nhiều cảm biến [19]. Ở đây vị trí của phương tiện được ước tính bằng hệ thống đo bằng hình ảnh sử dụng phương trình dưới dạng ước lượng cực đại hàm xác suất trong phương pháp lọc Kalman mở rộng nếu sai số nhỏ (thấp hơn ngưỡng cho phép)

$$\hat{P}_f(t) = P(t) + \Sigma_f(t)\Sigma_{gps}^{-1}(P_{gps}(t) - P(t)) \quad (2.35)$$

$$\Sigma_f(t) = (\Sigma_p(t)^{-1} + \Sigma_{gps}^{-1})^{-1} \quad (2.36)$$

Với $\hat{P}_f(t)$ và $\Sigma_f(t)$ là tương ứng với vị trí chính xác của phương tiện và sai số của ma trận sau khi kết hợp. $\hat{P}_f(t)$ và $\Sigma_f(t)$ được thay thế bằng $\hat{P}_f(t)$ và $\Sigma_f(t)$.

Chương 3

THỰC NGHIỆM GIẢI THUẬT

Trong một ngày những thực nghiệm sẽ được tiến hành để đánh giá tính hiệu quả của phương pháp đề xuất. Thực hiện được tiến hành trong một khoảng di chuyển dài trên một vùng đô thị rộng lớn có diện tích 0.7×0.7 km. Đoạn video và GPS được thu nhận từ camera cầu có dạng hyperbol và thiết bị thu nhận GPS được gắn bên trên thiết bị như trong hình 12. Trong thực nghiệm này đoạn đường trên bản đồ được xem như là quỹ đạo thực được đánh dấu bằng những đường màu vàng đứt nét trong hình 14. Hai thực nghiệm được tiến hành trong nghiên cứu này, ở thí nghiệm thứ nhất là sự so sánh của quỹ đạo 3D với chuyển động giả định trên 2D. Thí nghiệm thứ hai là sự so sánh của hệ thống thị giác so với hệ thống thị giác kết hợp với GPS.



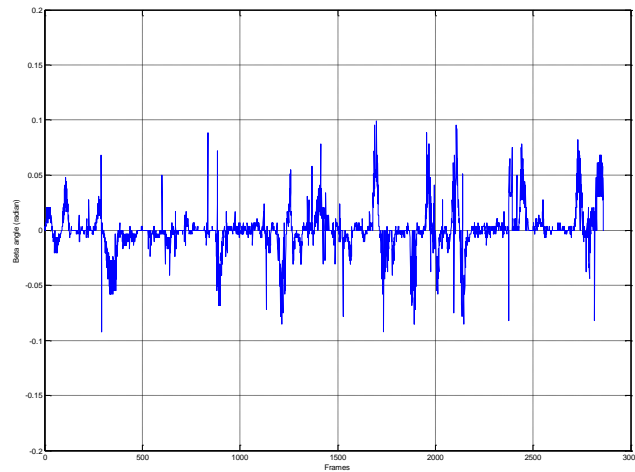
Hình 12: Xe gắn với nhiều cảm biến:
(a) cảm biến trên xe (b) Xe điện

Trong thí nghiệm thứ nhất, khoảng cách di chuyển của đối tượng là 1.62 km và 5716 khung hình được xử lý. Những hình ảnh từ camera cầu được trải ra dưới dạng panorama và hướng của đối tượng được xác định trước tại khung hình đầu tiên để tìm góc β , một vùng ảnh có kích cỡ 50x240 trên khung hình đầu tiên sẽ trượt lên trên khung hình thứ hai. Ngược lại với phương pháp đối sánh dùng Normalize Cross Correlation (NCC) hay là Sum of Squared Differences (SSD) thường yêu cầu khối lượng tính toán nhiều, ở đây phương pháp đối sánh chamfer được thực hiện vùng biên của vùng ảnh mẫu và ảnh panorama của khung hình thứ hai được tạo ra bằng phương pháp phát hiện biên kanny. để giảm những sai số, giá trị của góc β phải nhỏ hơn $\beta_{\max} = 15^{\circ}$. Điều này luôn luôn đúng trong thực tế. Thậm chí nếu góc của đường đi là 90° hay nhỏ hơn, thì góc này cũng thay đổi ở một hình dạng cung tròn bởi vì cấu trúc cơ khí và mô hình động học của đối tượng. Với những hạn chế này những sai số lớn về hướng gây ra bởi những vật mốc ở gần được loại bỏ. Hình 13 trình bày sự thay đổi của góc này theo những khung hình liên tục

Đồng thời trong bước xử lý song song với nó, những đặc trưng SIFT cũng được trích ra từ những khung hình liên tục. Những tính toán của góc alpha từ ma trận thiết hiệu được kết hợp với thuật toán RANSAC để loại bỏ những đặc trưng sai. Trong thí nghiệm này tốc độ di chuyển của đối tượng là hằng số và bằng 10 km/h. Vector dịch chuyển tịnh tiến được tính toán theo thời gian bằng cách nhân với tốc độ của nó. Cuối cùng những sai số trong việc so sánh quỹ đạo 3D và giả định chuyển động trong mặt phẳng 2D được trình bày trong bảng 2.

Bảng 2: So sánh chuyển động trên quỹ đạo 3D và 2D

Phương pháp	Giả thuyết 2D	Quỹ đạo 3D
Khoảng cách	1758	1557
Sai số	10.85%	3.89%



Hình 13: Giá trị của β so với số khung ảnh

Trong thí nghiệm thứ hai là sự so sánh của hệ thống thị giác hoạt động độc lập và hệ thống thị giác kết hợp với GPS. Như được giải thích ở phần trên, quỹ đạo của hệ thống thị giác thường phải gánh chịu những sai số tích lũy, vì vậy quỹ đạo này sẽ bị phân kỳ so với quỹ đạo thật khi khoảng cách di chuyển xa như trình bày trong hình 14. Vị trí chính xác của GPS được chọn lựa từ cách đo độ lệch. Nếu x_{ch} < 0.3 m và y_{ch} < 0.3 m thì nó được xem như là vị trí đúng. Sau khi chọn lựa được vị trí đúng từ GPS những vị trí đối tượng bị phân kỳ trên bản đồ sẽ được hiệu chỉnh lại. Quá trình này sẽ được tương tác sau mỗi 50 khung hình. So sánh những sai số này được trình bày trong bảng 3.

Bảng 3. So sánh sai số lớn nhất với giá trị quỹ đạo thực

Phương pháp	VO	VO + GPS
Độ sai số lớn nhất	23.35	0.53

Khi đối tượng di chuyển ở dưới các đường hầm hay đô thị mật độ xây dựng cao, tín hiệu GPS sẽ bị mất hay bị nhảy bước. Khi đó quỹ đạo sẽ được xây dựng chỉ dựa trên hệ thống thị giác. Việc đánh giá hệ thống GPS và quá trình hiệu chỉnh được khôi phục lại một khi những tín hiệu này xuất hiện trở lại. Đường quỹ đạo, đường quỹ đạo thực, quỹ đạo dùng hệ thống thị giác và GPS kết hợp với hệ thống thị giác được trình bày trong hình 14.



Hình 14. Kết quả định vị

Chương 4

KẾT LUẬN

Việc định vị cho thiết bị di chuyển ngoài trời dùng giải thuật 1 điểm kết hợp với một vùng đường biên của camera cầu đa hướng và hiệu chỉnh bằng GPS trong khoảng di chuyển xa đã được trình bày trong nghiên cứu này. Những ưu điểm đã được chỉ ra từ những phân tích và thực nghiệm. Đầu tiên, phương pháp đề xuất này đã cải thiện giả thuyết chuyển động là mặt phẳng trong hầu hết những nghiên cứu trước đây về định vị đối tượng di chuyển. Phương pháp này cũng đã tận dụng mô hình di chuyển tựa như xe hơi. Những ràng buộc trong những khung hình liên tục thì được giảm xuống còn 1 điểm kết hợp với 1 vùng đối sánh đường biên. Điều đó đơn giản hơn so với những phương pháp truyền thống. Sử dụng ít nhất 5 điểm tương đồng. Thứ hai vấn đề sai số tích lũy thì được loại bỏ trong khoảng di chuyển xa, điều mà có thể dẫn tới sự phân kỳ của quỹ đạo khi so sánh với quỹ đạo thực. Nói một cách khác, quỹ đạo hay vị trí của đối tượng đôi khi không đúng ở những vị trí cục bộ nhưng quỹ đạo toàn cục thường không bị sai. Hệ thống thị giác có thể hoạt động độc lập nếu tín hiệu GPS bị mất hoặc dao động trong một khoảng ngắn. Hơn nữa, hệ thống đo lường dùng thị giác cũng có thể làm việc một cách chính xác trong môi trường có cấu trúc phức tạp hay những điều kiện ánh sáng khác nhau. Tính bền vững và tốc độ xử lý của giải thuật này có thể được áp dụng trong các hệ thống thực. Những công việc trong tương lai sẽ tập trung vào việc nâng cao thời gian xử lý cho các hệ thống thời gian thực để tạo ra kết quả định vị chính xác hơn.

PHỤ LỤC A

Mối liên hệ giữa góc β và ψ

Từ phân tích hình học trên hình 15, điểm K là ảnh của P bằng cách xoay quanh OI. Do đó $OP=OK$. Góc OIK và OIH được tính như sau:

$$\tan \beta = \frac{OK}{OI} \quad (\text{A.1})$$

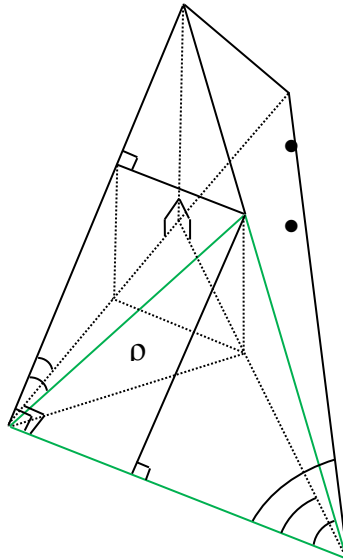
$$\tan \psi = \frac{OH}{OI} \quad (\text{A.2})$$

Do đó

$$\frac{\tan \beta}{\tan \psi} = \frac{OK}{OH} = \frac{OP}{OH} = \frac{1}{\cos \alpha} \quad (\text{A.3})$$

Bởi vì β và ψ rất bé nên có thể xấp xỉ $\tan \beta = \beta$, $\tan \psi = \psi$, do đó:

$$\frac{\beta}{\psi} = \frac{1}{\cos \alpha} \quad (\text{A.4})$$



Hình 15: Phân tích hình học của mô hình chuyển động cung tròn

TÀI LIỆU THAM KHẢO

- [1] D. Nistér, O. Naroditsky, and J. Bergen. “Visual odometer,” Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. I652–I659, 2004.
- [2] E. Royer, M. Lhuillier, M. Dhome, and J. M. Lavest, “Monocular vision for mobile robot localization and autonomous navigation,” International Journal of Computer Vision, vol. 74, no. 3, pp. 237–260, 2007.
- [3] R. García-García, M. A. Sotelo, I. Parra, D. Fernández, J. E. Naranjo, and M. Gavilán, “3D visual odometry for road vehicles,” Journal of Intelligent and Robotic Systems, vol. 51, no. 1, pp. 113–134, 2008.
- [4] K. Konolige, M. Agrawal, and J. Solà, “Large-scale visual odometry for rough terrain,” Proceedings of the International Symposium on Robotics Research, 2007.
- [5] Scaramuzza, D., “1-Point-RANSAC Structure from Motion for Vehicle-Mounted Cameras by Exploiting Non-holonomic Constraints,” International Journal of Computer Vision, Vol. 95, pp. 74-85, 2011
- [6] T. Gandhi and M. Trivedi, “Parametric ego-motion estimation for vehicle surround analysis using an omnidirectional camera,” Machine Vision and Applications, vol. 16, no. 2, pp. 85–95, 2005.
- [7] J. P. Tardif, Y. Pavlidis, and K. Daniilidis, “Monocular visual odometry in urban environments using an omnidirectional camera,” Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2531–2538, 2008.
- [8] S. Y. Kim, K. S. Yoon, D. H. Lee, and M. H. Lee, “The Localization of a Mobile Robot Using a Pseudolite Ultrasonic System and a Dead Reckoning

-
- Integrated System,” International Journal of Control, Automation, and Systems, vol. 9, no. 2, pp.339-347, 2011
- [9] S. Sukkarieh, E. M. Nebot, and H. F. Durrant-Whyte, “A high integrity IMU/GPS navigation loop for autonomous land vehicle applications,” IEEE Transactions on Robotics and Automation, vol. 15, no. 3, pp. 572–578, 1999.
- [10] J.-H. Choi, Y.-W. Park, J.-B. Song, and I.-S. Kweon, “Localization using GPS and VISION aided INS with an Image Database and a Network of a Ground-based Reference Station in Outdoor Environments,” International Journal of Control, Automation, and Systems, vol. 9, no. 4, pp. 716-725, 2011.
- [11] Y.-J. Lee, B.-D. Yim, and J.-B. Song, “Mobile Robot Localization based on Effective Combination of Vision and Range Sensors,” International Journal of Control, Automation, and Systems, vol. 7, no. 1, pp.97-104, 2009
- [12] Lijun Wei, Cindy Cappelle, Yassine Ruichek, and Frédérick Zann, “GPS and Stereovision-Based Visual Odometry: Application to Urban Scene Mapping and Intelligent Vehicle Localization,” International Journal of Vehicular Technology, vol. 2011, Article ID 439074, 17 pages, 2011.
- [13] Trần Đại Nghĩa, “Định vị cho robot di động dùng bộ lọc particle”, Luận văn thạc sỹ, ĐH Sư phạm Kỹ thuật TP.HCM, 2010, người hướng dẫn, TS. Ngô Văn Thuyên.
- [14] Ngô Văn Thuyên, Nguyễn Đăng Phúc Nguyên, “Thuật toán định vị và lập kế hoạch đường đi toàn cục cho robot tự hành”, *Tạp chí Khoa học và Công nghệ các trường Đại học Kỹ thuật*, số 85, năm 2011, trang 63-68
- [15] Nguyễn Vũ Sỹ, “Tránh vật cản cho robot di động bằng giải thuật biểu đồ trường vector”, Luận văn thạc sỹ, ĐH Sư phạm Kỹ thuật TP.HCM, 2011, người hướng dẫn, TS. Ngô Văn Thuyên.

-
- [16] Nguyễn Hùng Đông, “Xây dựng robot di động và thư viện mã nguồn mở trên Player”, Luận văn thạc sỹ, ĐH Sư phạm Kỹ thuật TP.HCM, 2011, người hướng dẫn, TS. Ngô Văn Thuyên.
- [17] Nguyễn Hữu Nam, “Quy hoạch quỹ đạo cho robot di động dựa trên thị giác máy tính”, Luận văn thạc sỹ, ĐH Sư phạm Kỹ thuật TP.HCM, 2011, người hướng dẫn, TS. Ngô Văn Thuyên.
- [18] Lê Hoàng Anh, “Định vị cho robot di động dùng camera”, Luận văn thạc sỹ, ĐH Sư phạm Kỹ thuật TP.HCM, 2012, người hướng dẫn, TS. Ngô Văn Thuyên.
- [19] Thanh H. Nguyen, Hung T. Nguyen, “A Bayesian Recursive Algorithm for Freespace Estimation Using a Stereoscopic Camera System in an Autonomous Wheelchair,” *Scientific & Academic Publishing on American Journal of Biomedical Engineering*, Vol. 1, No. 1, pp. 44-54, 2011.
- [20] Thanh Hai Nguyen, Vo Tuan Kiet (2010) “Freespace Estimation in an Autonomous Wheelchair Using a Stereoscopic Cameras System,” *proceedings of the 32rd IEEE/ EMBS Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Buenos Aires, Argentina, August 31 – September 4, 2010.

