

Nghiên cứu so sánh các thuật toán xử lý ảnh tính độ sâu ảnh stereo ứng dụng trong hệ thống camera thị giác

Bạch Ngọc Minh*

Trung tâm Phát triển Đại học Quốc gia Hà Nội, 144 Xuân Thủy, Hà Nội, Việt Nam

Nhận ngày 16 tháng 7 năm 2014

Chỉnh sửa ngày 18 tháng 8 năm 2014; Chấp nhận đăng ngày 22 tháng 9 năm 2014

Tóm tắt: Có ba phương pháp cơ bản để so sánh sự tương hợp giữa hai ảnh là SAD, SSD và NCC. Phương pháp NCC sử dụng sự tương quan chéo giữa các điểm ảnh nên các phép toán rất phức tạp. Chúng bao gồm phép toán bình phương, tổng, tích các tổng, khai căn nên việc thực hiện chắc chắn rất tốn thời gian. Còn hai phương pháp SAD và SSD có độ phức tạp của thuật toán không khác biệt quá lớn nên chúng ta cần làm một thực nghiệm để kiểm chứng chất lượng bản đồ độ sâu và tốc độ thực hiện. Từ đó chọn ra một phương pháp thích hợp hơn

So sánh thời gian chạy giữa hai giải thuật SAD và SSD cho thấy việc lựa chọn các thuật toán SAD là hợp lý hơn, bởi vì các thuật toán SSD cần gần gấp đôi thời gian chạy. Phương pháp SAD có thể giải quyết vấn đề đối với việc tìm độ sâu trong ảnh 3D. Một điểm ở không gian lấy ảnh ở hai ống kính máy tại hai vị trí khác nhau. Chính vì thế nên trước khi tìm được độ sâu của điểm ảnh, ta cần tìm được các điểm ảnh tương ứng đã. Phương pháp SAD là phương pháp đơn giản nhất để tìm được điểm ảnh đó.

Phương pháp SAD lại có nhiều cách tính toán, tính theo phương pháp SAD truyền thống, phương pháp SAD điều chỉnh (dựa vào biên ảnh), phương pháp SAD phân đoạn (phân vùng ảnh). Ở đây ta lại lựa chọn giữa 2 phương pháp SAD điều chỉnh và SAD phân đoạn giảm 20% so với SAD truyền thống.

Tiếp theo chúng ta xem xét các thuật toán SAD phân vùng lại được tạo ra bởi sự kết hợp của hai kỹ thuật: Belief Propagation và các thuật toán phân đoạn Mean Shift. Cách tiếp cận này kết hợp những ưu điểm của cả hai phương pháp phân đoạn. Các thuật toán Mean Shift cho kết quả nhanh chóng và Belief Propagation thực hiện phân đoạn rất chính xác.

Từ khóa: Phân vùng ảnh, Belief Propagation, Mean Shift, SAD, hình ảnh 3D.

1. Giới thiệu

Một xu hướng phát triển mới là rô bốt sử dụng camera thị giác để quan sát như mắt người. Mặc dù việc chiết xuất dữ liệu từ ảnh thị

giác là khó khăn hơn là dùng cảm biến nhưng bù lại thì cách thức này cung cấp cho rô bốt và con người những dữ liệu trực quan sinh động và đầy đủ.

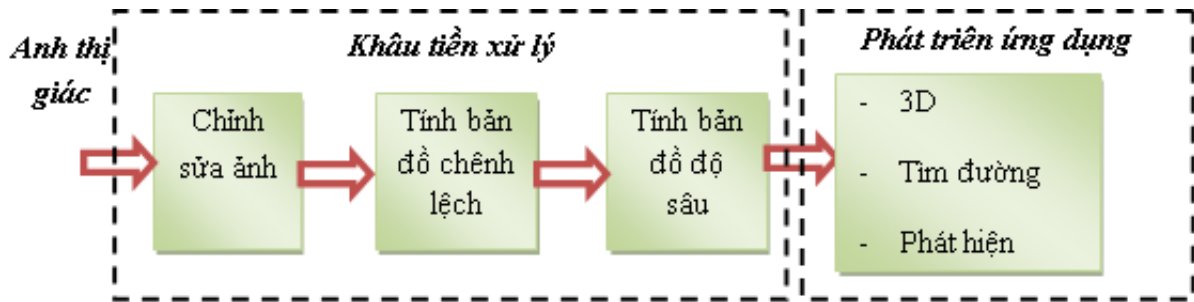
Khi xử lý tín hiệu ảnh stereo thị giác thường có hai khâu (Hình 1).

*ĐT.: 84-913550789

Email: minhbn@vnu.edu.vn.

* Khâu đầu tiên là khâu tiền xử lý có nhiệm vụ lấy ra các dữ liệu có chứa thông tin hữu ích cho việc phát triển các ứng dụng thị giác stereo thời gian thực.

* Khâu thứ hai các ứng dụng thị giác như ứng dụng tái tạo môi trường 3D, tìm đường, phát hiện đối tượng...



Hình 1. Quá trình xử lý dữ liệu ảnh stereo thị giác.

Trong khâu tiền xử lý, có một đặc điểm chính là khối lượng dữ liệu thu thập bằng camera rất lớn (lớn hơn nhiều lần so với dữ liệu thu thập bằng cảm biến) cần được xử lý.

Mọi cách giải quyết mới đều phải được phát triển trên một nền tảng sẵn có. Ở phần tiếp, các kỹ thuật cơ bản của xử lý ảnh 3D sẽ được đưa ra để làm cơ sở phát triển cho các thuật toán về sau. Bước đầu tiên, chúng ta sẽ tìm hiểu về ảnh 3D và máy quay 3D, nền tảng cơ sở đầu

tiên của hệ thống. Sau đó, các kỹ thuật xử lý để trích ra nội dung của ảnh sẽ được bàn đến. Các thuật toán cơ bản để dựa vào đó mà có các nghiên cứu hoàn thiện hơn.

1.1. Ảnh 3D: Là ảnh mà bản thân nó có chứa nội dung về chiều sâu. Các điểm ảnh ngoài các giá trị về màu sắc, tọa độ theo hai trục cơ bản còn có giá trị về chiều sâu. Để tạo nên được tầm ảnh 3D, người ta thực hiện đồng thời hai tầm ảnh về cùng một nội dung nhưng khác hướng.



Hình 2. Một cảnh trong phim 3D.

(Hình 2) thể hiện một tấm ảnh 3D được theo dõi bằng mắt kính chuyên dụng. Ảnh trên được cấu tạo từ hai tấm ảnh khác nhau được trộn bởi 2 dải màu khác nhau. Khi nhìn ảnh trên bằng kính mắt chuyên dụng, mỗi mắt của người xem sẽ thấy một ảnh. Sự chênh lệch vị trí trong 2 tấm ảnh nhìn thấy sẽ tạo nên cảm giác về độ sâu cho người xem. Thực tế, khi tách hai tấm ảnh của một ảnh 3D, chúng ta sẽ có thể nhận được những bức ảnh tương tự như trong (Hình 3).

Ảnh 3D có nhiều ứng dụng hơn là để giải trí. Nó có khả năng lưu trữ độ sâu điểm ảnh nên cũng được sử dụng cho các ứng dụng yêu cầu độ sâu điểm ảnh. Nói cách khác, từ một tấm ảnh 3D, chúng ta có thể mô tả lại không gian hoặc lấy những thông tin hữu ích từ độ sâu điểm ảnh [1]. Các đối tượng 3D cũng có thể được trích ra từ một tấm ảnh 3D.



a



b

Hình 3. Ảnh 3D sau khi tách ra. a: Ảnh nhìn thấy ở mắt trái; b: Ảnh nhìn thấy ở mắt phải

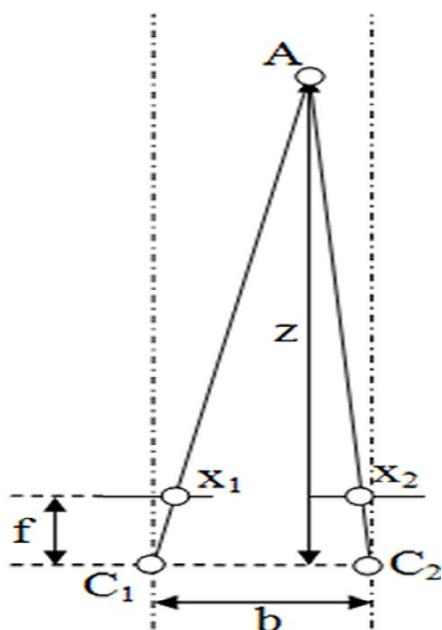
1.2. Stereo camera: Là thiết bị thu ảnh 3D hoặc thu phim 3D. Về cơ bản, stereo camera cũng giống những máy ảnh hoặc máy quay chuyên dụng khác, nhưng nó có hai ống kính mắc song song (

Hình 4)[2]. Hai ống kính của một stereo-camera được sử dụng để lấy hai hình ảnh trong một ảnh 3D.



Hình 4. Hệ thống Stereo camera Bumblebee [3].

(Hình 5). thể hiện hoạt động của một stereo camera [3]. Các chi tiết sẽ được thu vào 2 tấm ảnh thông qua 2 ống kính của camera cùng một lúc. Lúc này, tọa độ của điểm ảnh thu được trên hai tấm sẽ có một độ lệch nhất định, ta gọi độ lệch đó là Δx .



Hình 5. Hoạt động của stereo camera.

Ta có:

$$\Delta x = x_1 - x_2 \quad (1-1)$$

Và khoảng cách từ điểm A đến camera sẽ được tính:

$$z = \frac{f \times b}{\Delta x} \quad (1-2)$$

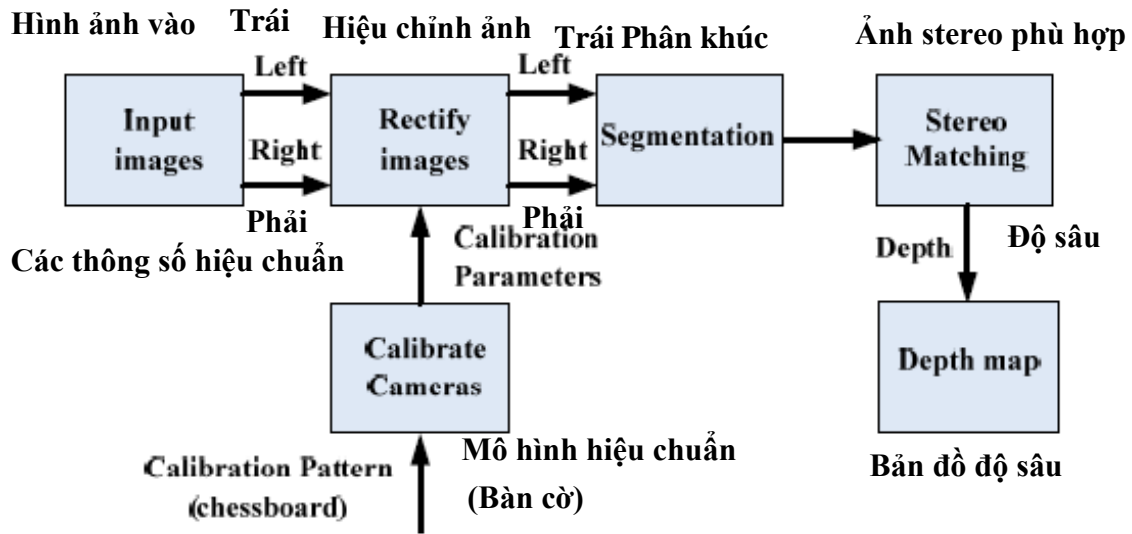
Như vậy, từ các tọa độ ảnh thu được của điểm A trên hai tấm ảnh trái phải, ta sẽ thu được khoảng cách từ camera tới điểm A. Khoảng cách này còn được gọi là độ sâu điểm ảnh ký hiệu bằng z.

1.3. Tính toán bản đồ chênh lệch: Là một trong những vấn đề quan trọng trong thị giác máy tính 3D. Một số lượng lớn các thuật toán đã được đề xuất để giải quyết vấn đề này [4],[5]. Một trong những phương pháp tương đối mới là Cải thiện tính toán Bản đồ độ sâu từ hình ảnh stereo theo phương pháp lai.

Đối với các cặp hình ảnh stereo đã được chỉnh sửa epipolar, mỗi điểm trong hình ảnh bên trái nằm trên đường nằm ngang (đường epipolar) có thể có điểm ảnh tương ứng trong hình ảnh bên phải. Cách tiếp cận này được sử dụng để làm giảm không gian tìm kiếm chiều sâu bản đồ thuật toán tính toán. Chiều sâu của một điểm ảnh là khoảng cách điểm không gian tương ứng tới trung tâm máy ảnh. Để ước tính bản đồ độ sâu và phát hiện các đối tượng 3D, các điểm ảnh tương ứng trong những ảnh trái và ảnh phải cần được phát hiện. Thuật toán này bao gồm các giai đoạn sau đây (Hình 6) :

1. Thu thập hình ảnh,
2. Hình học Epipolar và cải thiện hình ảnh,
3. Phân đoạn,
4. Thuật toán phù hợp stereo,
5. Ước lượng bản đồ độ sâu.

Đầu tiên, biến dạng xuyên tâm và tiếp tuyến của ống kính được loại bỏ bằng cách hiệu chỉnh máy ảnh bởi các thông số bên trong và bên ngoài máy ảnh. Để làm được điều này cần có sự hiểu biết về các thông số máy ảnh với mục đích khắc phục cả hai hình ảnh. Sau khi hiệu chỉnh, hình ảnh được tách ra thành các khu vực bằng cách sử dụng các thuật toán phân chia lai được đề xuất. Cuối cùng, thuật toán phù hợp stereo được áp dụng trên các hình ảnh phân đoạn trái và phải với mục đích để tìm tất cả các tương quan (điểm phù hợp) và gán chiều sâu cho từng phân đoạn. Đầu ra của thuật toán phù hợp stereo là các bản đồ độ sâu.



Hình 6. Thuật toán SAD phân vùng [6].

Trong phần tiếp theo, mục 2.1 trình bày về tổng quan về hiệu chuẩn máy ảnh cơ bản. Hình ảnh cải tiến dựa trên phân đoạn được mô tả trong mục 2.3. Phương pháp phân đoạn Mean Shift và Belief Propagation được trình bày trong mục 6 và mục 7. Tại mục 8 trình bày về thuật toán SAD phân vùng lại. Cuối cùng là kết quả thực nghiệm và kết luận được trình bày ở mục 9 và mục 10.

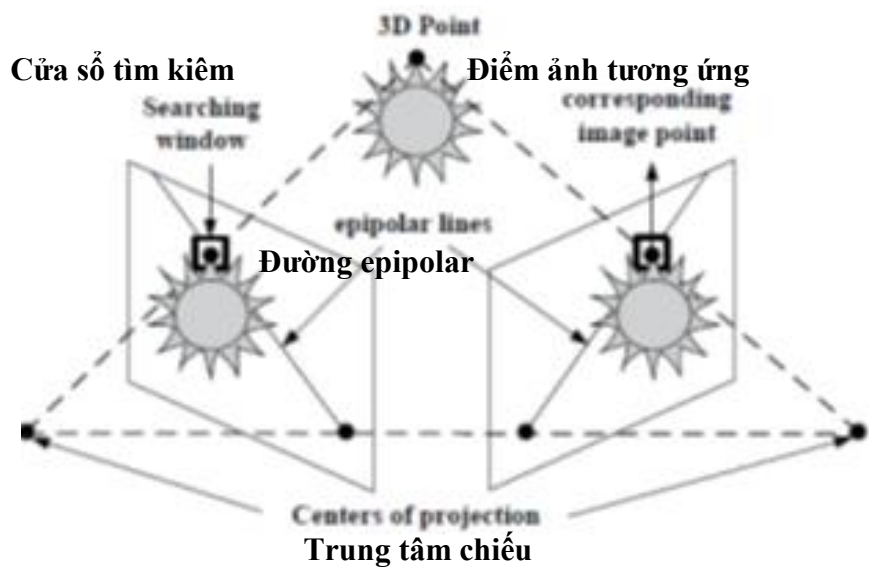
2. Cải thiện hình ảnh

2.1. Hiệu chuẩn máy ảnh: Để áp dụng các kỹ thuật khác nhau cho ảnh stereo với một mức độ hợp lý chính xác, điều quan trọng là để hiệu chỉnh hệ thống camera. Nó là một quá trình tìm kiếm các thông số bên trong và bên ngoài của máy ảnh. Các phương pháp hiệu chuẩn cổ điển dựa trên các mẫu hiệu chuẩn đặc biệt chuẩn bị, các đối tượng với kích thước và vị trí được biết đến trong một hệ tọa độ nhất định. Sau đó, các tính năng, chẳng hạn như góc và đường, được chiết xuất từ một hình ảnh của mô hình chuẩn. Đối tượng với các tính năng có ý nghĩa thường

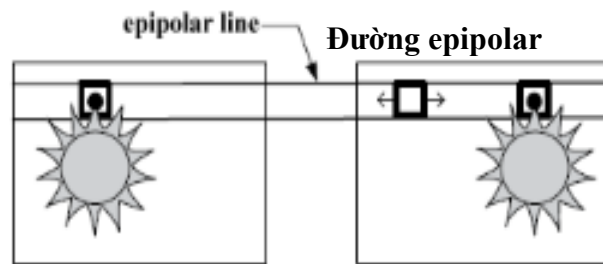
được chọn để hiệu chỉnh được một cách rõ ràng các vị trí của chúng. Người ta thường dùng một bàn cờ đơn giản có thể đạt được mục đích này.

2.2. Chỉnh sửa hình ảnh: Chỉnh sửa hình ảnh (Hình 7) là cần thiết để giảm độ phức tạp tính toán điểm ảnh tương ứng trong hình ảnh trái và bên phải. Mục đích của việc cải thiện hình ảnh là để tìm đường epipolar của hai hình ảnh theo liên kết chiều ngang. Điều này có thể được thực hiện bằng cách sử dụng các biến đổi tuyến tính xoay, dịch và nghiêng hình ảnh. Các thông số nội bộ camera và thông tin về định hướng và vị trí máy ảnh được sử dụng trong các phép biến đổi.

Các điểm 3D được phản chiếu thành các điểm trong hình ảnh stereo trái và phải. Sau khi cải thiện theo (Hình 8), các đường epipolar của hai điểm phản chiếu là song song theo chiều ngang và liên kết dọc theo mặt phẳng hình ảnh mới. Các điểm được nằm trên cùng một đường dây epipolar [7]. Do đó vấn đề phù hợp stereo được giảm xuống tìm kiếm một chiều dọc theo các đường ngang, thay vì tìm kiếm hai chiều như nó được hiển thị trong (Hình 8) [8],[9].



Hình 7. Ảnh stereo trước khi chỉnh sửa.



Hình 8. Ảnh stereo sau khi chỉnh sửa.

2.3. Phân đoạn hình ảnh màu: Trong phần này, hai phương pháp phân chia màu sắc được mô tả: Belief Propagation, Mean Shift. Mục tiêu của việc phân vùng ảnh là chia toàn bộ hình ảnh thành một tập hợp các phân đoạn bao gồm các hình ảnh. Các phân đoạn cuối cùng phải đáp ứng đầy đủ bốn điều kiện sau [10]:

1. $\bigcup_{i=1}^n R_i = R$
2. Với mọi i và j , $i \neq j$, tồn tại $R_i \cap R_j = \emptyset$
3. Với $i = 1, 2, \dots, n$ thì bắt buộc $P(R_i) = TRUE$

4. Với mọi $i \neq j$, tồn tại $P(R_i \cup R_j) = FALSE$

trong đó R đại diện cho toàn bộ hình ảnh, R_i ($i = 1, 2, \dots, n$) là các phân đoạn không rỗng, không chung nhau của R , $P(R_i)$ là thuộc tính giống nhau của tất cả các yếu tố trong R_i và 0 đại diện cho một tập rỗng. Tổng kết công tác phân đoạn phải bao gồm tất cả các điểm ảnh trong hình ảnh. Đây là điều kiện tiên quyết. Điều kiện thứ hai thể hiện các phân đoạn khác nhau không chồng chéo nhau. Điều kiện thứ ba thể hiện rằng các điểm ảnh trong cùng phân

đoạn phải có chung thuộc tính. Các thuộc tính phải khác nhau là điều kiện cuối cùng.

Có một số thuật toán đã được đề xuất để giải quyết vấn đề này. Tuy nhiên, tất cả các thuật toán này đều sử dụng hàm chi phí kết hợp để thiết lập sự tương ứng giữa hai điểm ảnh. Có

ba phương pháp phổ biến nhất là phương pháp dựa trên sự khác biệt cường độ tuyệt đối (AD), phương pháp dựa trên sự khác biệt cường độ bình phương (SD) và phương pháp dựa trên sự tương quan chéo chuẩn (NCC).

2.4. Các phương pháp thị giác stereo chính được thể hiện về mặt toán học như sau:

1- Phương pháp tổng sự khác biệt tuyệt đối (SAD - Sum of Absolute Differences):

$$SAD(x, y, d) = \sum_{x,y \in W} |I_l(x, y) - I_r(x, y - d)| \quad (2-3)$$

Theo công thức (2-3), việc tính giá trị chênh lệch của điểm (x,y) được thực hiện bằng phép tính hiệu trong cửa sổ W. Do đó thuật toán đơn giản. Cửa sổ W và phạm vi chênh lệch d càng lớn thì số phép tính tăng. Thông thường có thể

thực tính SAD với cửa sổ 3x3, hoặc 5x5 hoặc 7x7 hoặc 11x11. Phạm vi chênh lệch d thường chọn dưới 120 với các giá trị tham khảo như 16, 50, 128. [11-13].

2- Phương pháp tổng bình phương khác biệt (SSD – Sum of Squared Differences):

$$SSD(x, y, d) = \sum_{x,y \in W} (I_l(x, y) - I_r(x, y - d))^2 \quad (2-4)$$

Phương pháp này tính tổng của các bình phương của các hiệu nên xuất hiện thêm phép

nhân trước khi tính tổng. Do đó độ phức tạp tăng lên đáng kể [19].

3- Phương pháp dựa trên sự tương quan chéo chuẩn (NCC)

$$NCC(x, y, d) = \frac{\sum_{x,y \in W} I_l(x, y) \cdot I_r(x, y - d)}{\sqrt{\sum_{x,y \in W} I_l^2(x, y) \cdot \sum_{x,y \in W} I_r^2(x, y - d)}} \quad (2-5)$$

Trong đó I_l , I_r là những giá trị cường độ trong hình ảnh trái và phải, (x, y) là tọa độ của điểm ảnh, d là giá trị sai lệch được xem xét và W là cửa sổ phức hợp. Việc lựa chọn các giá trị sai lệch thích hợp cho mỗi điểm ảnh được thực hiện sau đó.

Phương pháp này có độ phức tạp cao nhất vì xuất hiện các phép toán tiêu tốn nhiều tài nguyên của máy tính như phép bình phương, khai căn. Do đó việc ứng dụng của thuật toán này trong các hệ thống thời gian thực là không khả thi.

$$D(x, y) = \arg \min SAD(x, y, d) \quad (2-6)$$

tức là, cho mỗi điểm ảnh (x, y) và cho giá trị không đổi có sự khác nhau d, chi phí tối thiểu được lựa chọn. Phương trình (2-6) được dùng cho phương pháp SAD. Tuy nhiên, trong nhiều trường hợp lựa chọn khác biệt là một quá trình lặp đi lặp lại, vì độ chênh lệch của mỗi điểm ảnh là tùy thuộc vào sự chênh lệch của các điểm ảnh lân cận [14].

Các thuật toán phân vùng lai là một sự kết hợp với phương pháp SAD được áp dụng phù hợp với hình ảnh stereo để tinh chỉnh các bản đồ độ sâu cuối cùng. Cả hai, các thuật toán phù hợp với âm thanh stereo dựa trên phương pháp mà không SAD phân đoạn và cách tiếp cận lai

đề xuất dựa trên phương pháp phân khúc hybrid là hiệu quả. Tập hình ảnh đã được tính toán trên 2,8 GHz bộ xử lý Intel Core i5 với 4 GB bộ nhớ DDR3 và kit phát triển FPGA DE2 để mô phỏng.

Kết quả thực nghiệm SAD

Độ phân giải	Cửa sổ	Thời gian(s)
400x266	3x3	46
400x266	5x5	58.5
400x266	9x9	105

Kết quả thực nghiệm SSD

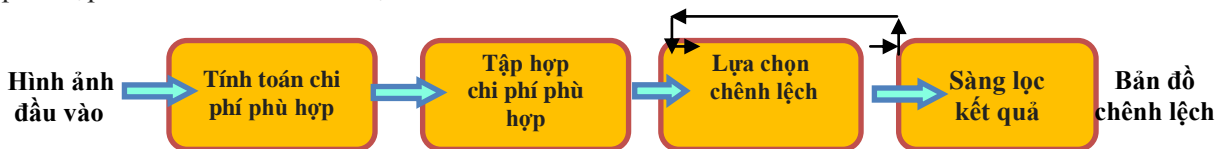
Độ phân giải	Cửa sổ	Thời gian(s)
400x266	3x3	106.5
400x266	5x5	335
400x266	9x9	891

Thời gian tính toán cho các thuật toán SAD là khoảng 105 giây và cho phương pháp phân vùng lai 29 giây, tương ứng. Các thuật toán phù hợp với hình ảnh stereo dựa trên SAD lại

có vẻ là thuật toán hiệu quả hơn để sản xuất bản đồ chênh lệch sạch hơn với các khu vực đồng nhất. Mặt khác, các thuật toán phù hợp với hình ảnh stereo dựa trên phương pháp SAD sản xuất một bản đồ chiều sâu rõ ràng về hiện trường. Hơn nữa, nó tạo ra một chất lượng cao hơn và ít lỗi gây ra bởi tắc phân đoạn hình ảnh.

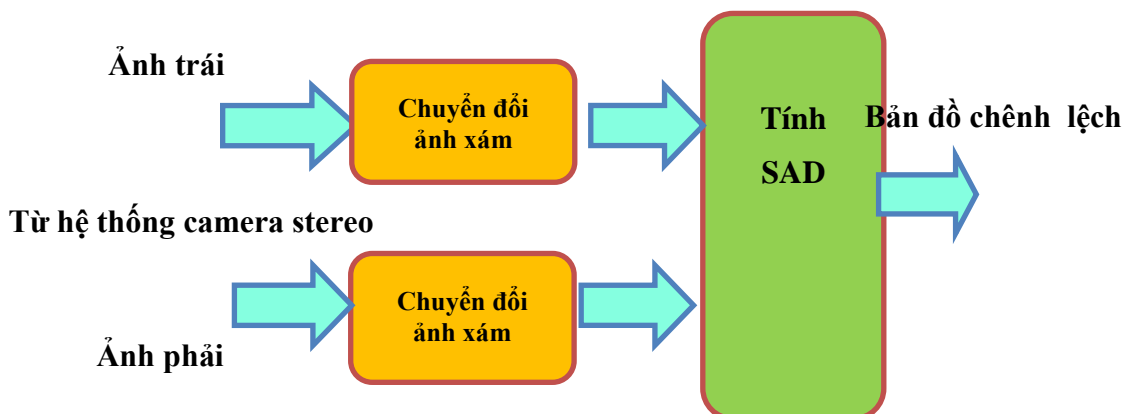
Nó được áp dụng cho các hình ảnh stereo chụp bởi máy ảnh lập thể của chúng tôi. Bộ dữ liệu thử nghiệm của chúng tôi bao gồm 25 cặp hình ảnh stereo thực sự trong màu xám quy mô với kích thước 800x600 pixels. (Hình 28)

Kết quả là, cần thiết có nhiều lần lặp lại để tìm ra các thiết lập tốt nhất của sự chênh lệch. Giai đoạn này tạo nên sự khác biệt giữa thuật toán địa phương với thuật toán toàn cục. Quá trình tính toán thường thêm một bước nữa để kết quả tốt hơn gọi là khâu lọc (Hình 9).



Hình 9. Sơ đồ khối tổng quát của một thuật toán tương ứng stereo.

3. Phương pháp SAD truyền thống:



Hình 10. Sơ đồ khối tổng quát của phương pháp SAD truyền thống.

Như đã phân tích ở phần 2.4, phương pháp SAD là phương pháp để nhận diện rất thuận tiện. Song, bên cạnh ưu điểm của phương pháp này, nó còn có một số khuyết điểm về thời gian đáp ứng và xác suất lỗi.

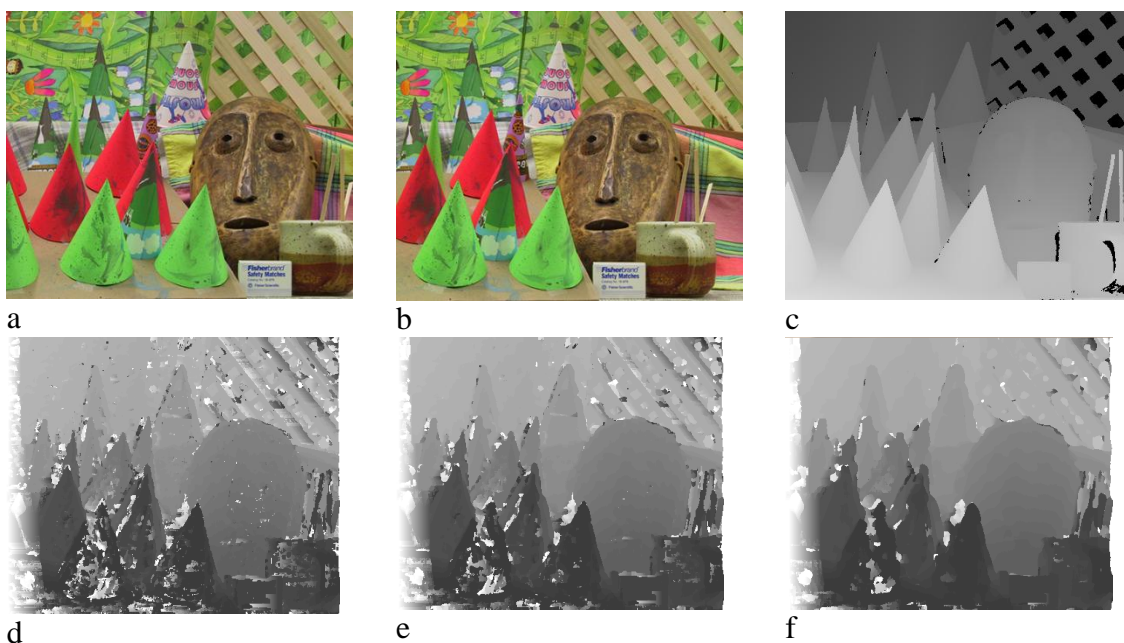
Công thức (3-7) thể hiện tốc độ của thuật toán này áp dụng vào việc tính Bản đồ chênh lệch.

$$t = S \times W \times L \quad (3-7)$$

Trong công thức này, t là thời gian chạy chương trình, S là số điểm cần tính, W là kích

thước cửa sổ tính toán và L là khoảng cách yêu cầu tính toán. Với công thức trên, giả sử cứ mỗi 1.5 triệu phép tính hàm chạy trong 1 giây, tấm ảnh xử lý với kích thước 640x480 pixel, kích thước cửa sổ là 5x5 pixel và khoảng cách thực hiện là 50 pixel.

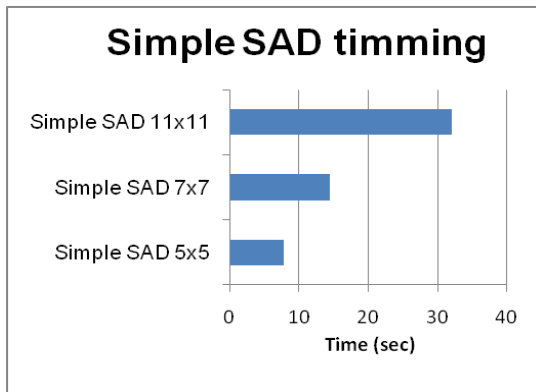
Khi đó, thời gian phép toán là 25.6 giây. Đây là một khoảng thời gian lớn, hay nói cách khác là tốc độ tính toán của phép toán là khá chậm.



Hình 11. Các kết quả tính SAD thông thường. a) Ảnh trái gốc, b) Ảnh phải gốc, c) Bản đồ chênh lệch mong muốn, d) Bản đồ chênh lệch tính bằng phương pháp SAD với cửa sổ 5x5, e) Bản đồ chênh lệch tính bằng phương pháp SAD với cửa sổ 7x7, f) Bản đồ chênh lệch tính bằng phương pháp SAD với cửa sổ 11x11.

Mặc dù tốc độ tính toán rất chậm, kết quả tính toán lại không chính xác hoàn toàn. Lỗi của phương pháp SAD rất thường xuyên xảy ra do nhầm lẫn trong quá trình tính toán. (Hình 11) thể hiện một số kết quả tính bản đồ sai lệch bằng phương pháp SAD với các cửa sổ 5x5, 7x7 và 11x11 thực hiện trên ảnh Teddy [14].

Nhận xét, ứng với các hình trên, lỗi cục bộ xảy ra nhiều ở những thuật toán sử dụng cửa sổ so sánh nhỏ. Tuy nhiên, hình dáng của các vật thể trong thuật toán lại được giữ gần nhất với hình dáng thực của nó. Sử dụng cửa sổ lớn hơn sẽ cho lỗi cục bộ ít hơn, nhưng hình dáng của các vật thể lại bị sai sót đáng kể so với thực tế. Hơn thế nữa, các lỗi lớn vẫn không thể giảm thiểu trong thuật toán này.



Hình 12. Tốc độ phép toán SAD với các kích thước cửa sổ khác nhau.

12) thể hiện thời gian của thuật toán đo dựa trên đầu vào là một tấm ảnh có kích thước 450x375. Như vậy, tốc độ của thuật toán khá bé, không thích hợp cho một hệ thống xử lý trên robot, kể cả khi đã loại bỏ tất cả các lỗi.

Lỗi của phương pháp SAD là do đâu? Phương pháp SAD dựa trên sự so sánh độ khác biệt giữa các cửa sổ ảnh. Theo như ý tưởng thuật toán thì ảnh của các điểm ở hai tấm ảnh là giống nhau nhất. Như vậy, kết quả tính toán của thuật toán này thực sự không hiệu quả cho hệ thống dẫn đường chạy trên robot. Nếu thực sự triển khai một hệ thống như vậy trên robot cứu

hộ thì sẽ phải tốn chi phí rất lớn về khả năng tính toán mà chưa chắc đã hiệu quả. Vì vậy, phương pháp này cần phải được cải thiện để có được kết quả tốt hơn và phù hợp hơn cho hệ thống thực.

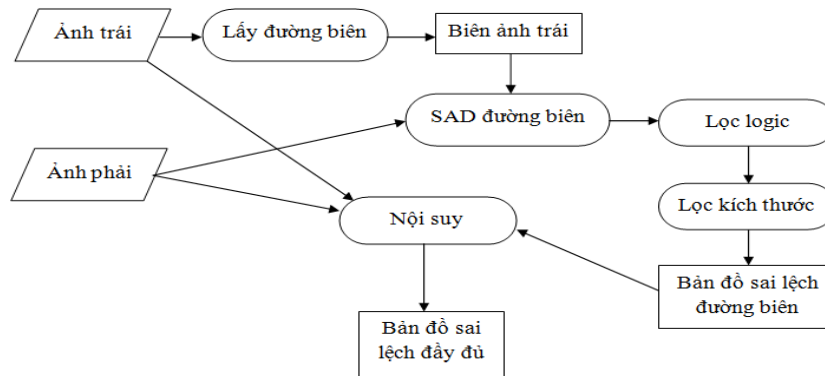
4. Phương pháp SAD điều chỉnh:

Sau khi phân tích những nhược điểm của phương pháp xác định bản đồ chênh lệch dựa trên phương pháp SAD truyền thống, tôi muốn đề xuất một phương pháp tốt hơn về chất lượng cũng như tốc độ.

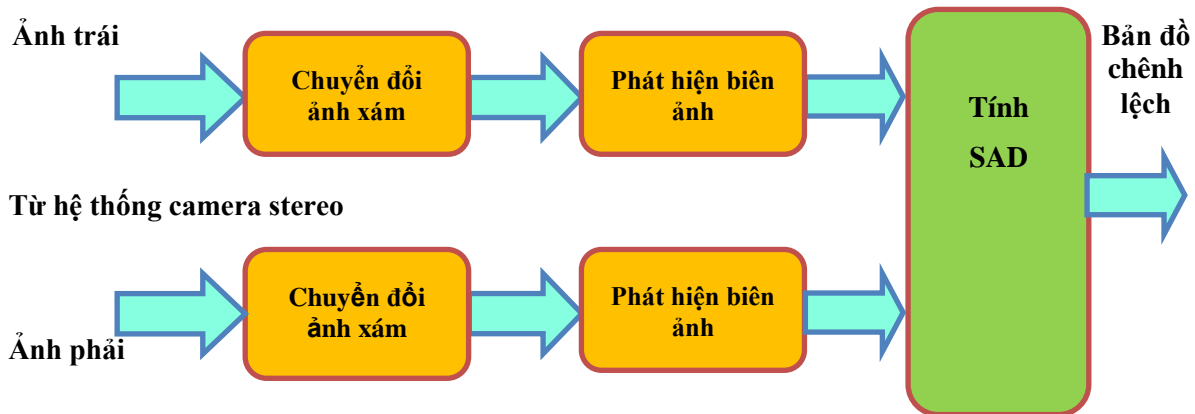
(Hình 13) thể hiện các bước của phương pháp mới nhằm lấy được bản đồ chênh lệch. Những mục tiếp theo của chương 3 sẽ thể hiện rõ từng bước của phương pháp này và ý nghĩa của chúng.

4.1. Phương pháp SAD đối với đường biên

Là phương pháp dựa trên phương pháp SAD nhưng chỉ áp dụng với những điểm thuộc đường biên. Với cách này, tôi kỳ vọng sẽ có được kết quả chính xác hơn với tốc độ tính toán cao hơn.

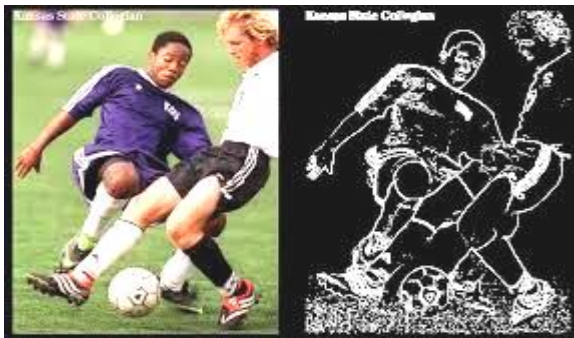


Hình 13. Phương pháp SAD điều chỉnh.



Hình 14. Sơ đồ khối tổng quát của phương pháp SAD điều chỉnh.

Giả sử có một cách nào đó, tôi có thể giảm số điểm trong tập hợp dữ liệu cần tính xuống còn 20% so với ban đầu. Như vậy, tốc độ của thuật toán có thể sẽ tăng lên đến 5 lần hay thời gian tính toán sẽ giảm xuống 5 lần. Thật vậy, theo phương trình (3 -7) nếu S giảm đi 5 lần và các giá trị khác vẫn giữ nguyên, ta sẽ có t giảm 5 lần. Như vậy, tôi sẽ tìm những điểm đáng tin cậy nhất để tính toán bản đồ chênh lệch nhằm giảm lỗi và tăng tốc độ.



Hình 15. Đường biên cũng là đường bao đối tượng.

Những điểm được chọn là những điểm mang tính phân biệt cao, hay nói cách khác là khó bị nhầm lẫn với những điểm khác. Khi chọn đường biên làm tập hợp dữ liệu để tính

toán, bước đầu ta cũng có thể tránh được sự nhầm lẫn gây ra bởi các điểm lân cận, lỗi nhiều nhất trong những phép toán trước.

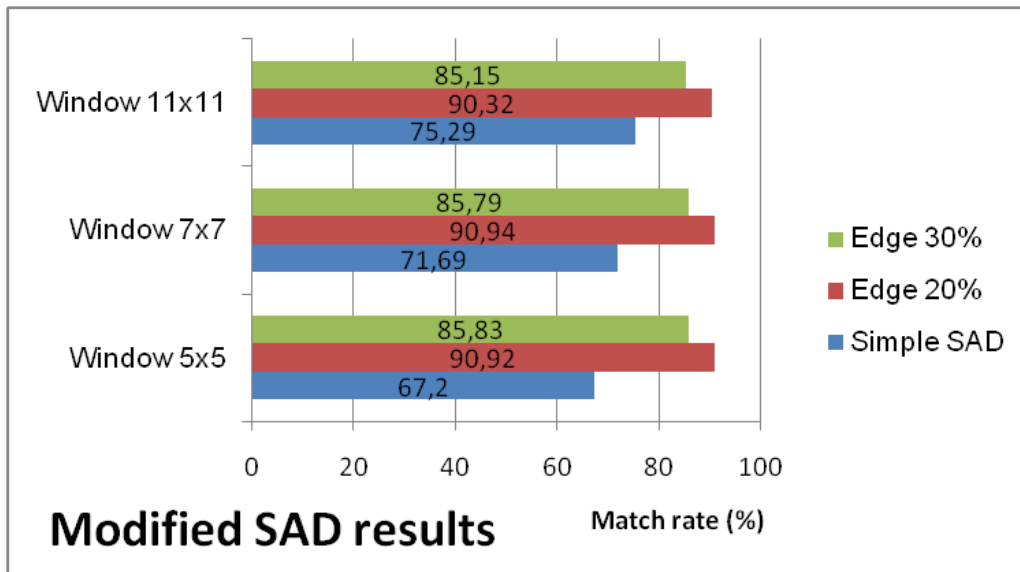
Đường biên còn giúp tạo ra kết quả nội suy chính xác hơn về sau. Như vậy, khi có được toàn bộ đường bao của đối tượng và độ sâu của tất cả các điểm trên nó, ta có thể suy ra độ sâu của những điểm còn lại.

Dưới những lý do trên, tôi chọn đường biên là những điểm sẽ được tính toán độ sâu trực tiếp. (Hình 16) là một sự so sánh kết quả tính bản đồ chênh lệch dựa trên đường biên là 20% bức ảnh, đường biên là 30% bức ảnh và toàn bộ bức ảnh. Với kết quả đã thể hiện trong (Hình 16), chúng ta rút ra một số nhận xét:

* Kết quả của thuật toán chính xác hơn đối với những đường biên. Số đường biên tính toán càng nhỏ, độ chính xác càng cao.

* Khi số điểm làm đường biên được tính toán nhỏ đến một ngưỡng nào đó thì kích thước cửa sổ tính toán không còn ảnh hưởng đến kết quả tính toán nữa.

Với nhận xét thứ 2, ta hiển nhiên có thể giảm kích thước cửa sổ lấy SAD đi mà vẫn kỳ vọng có được kết quả có độ chính xác cao.



Hình 16. Đánh giá kết quả của phương pháp SAD dựa trên đường biên.

4.2. Phương pháp trích chọn đường biên màu

Phương pháp trích chọn đường biên màu là một phương pháp cải tiến hơn so với phương pháp trích chọn đường biên dựa trên ảnh xám. Sở dĩ phải trích chọn đường biên màu vì đường biên ảnh xám không thể hiện được hết các đường bao của đối tượng. Các ảnh xám không thể thể hiện được hết các tính chất như đối với ảnh màu. (Hình 17) thể hiện sự khác biệt giữa đường biên được lấy từ ảnh màu và đường biên được lấy từ ảnh xám.

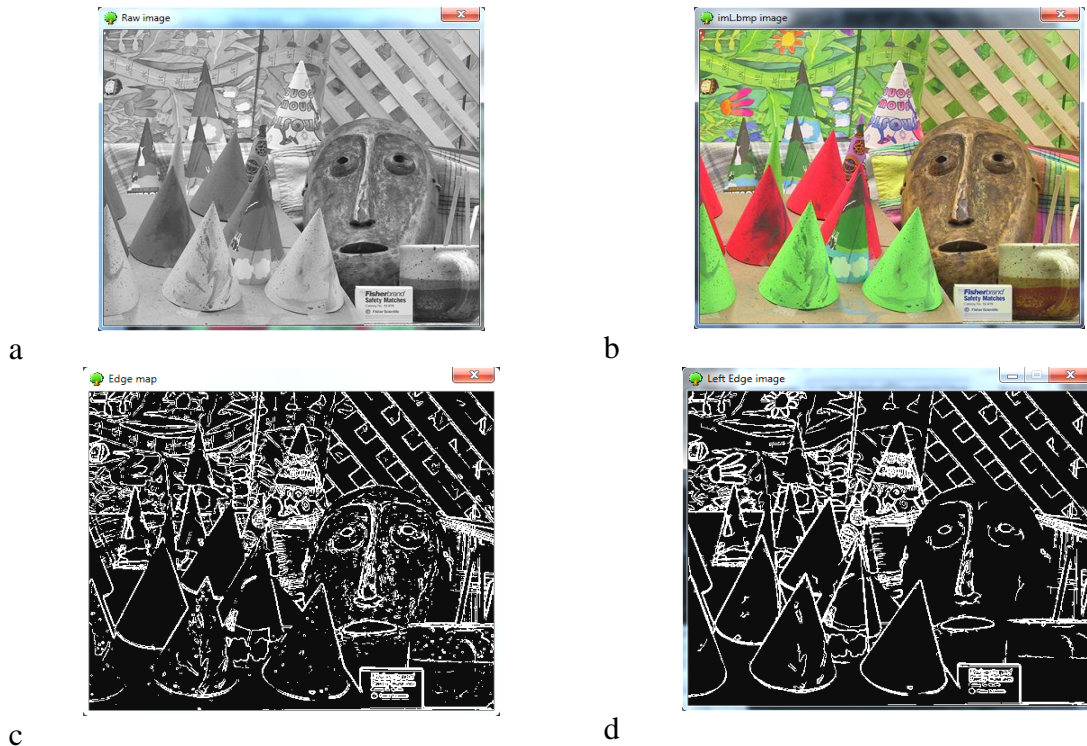
(Hình 18) thể hiện lý do của sự thiếu đường biên trong đường biên ảnh xám. Trong ảnh màu hoặc các mảng màu của ảnh gốc có những điểm có sự nhảy bậc của màu sắc hoặc đường biên, nhưng trong ảnh xám không có. Như vậy, thực hiện bất các đường biên bằng các mảng màu cơ bản, ta sẽ có các đường biên kín hơn và đặc trưng hơn. Sau đây sẽ là các bước để xây dựng giải thuật tìm đường biên dựa trên các mảng màu cơ bản:

* Bước 1: Ta coi 3 mảng màu giống như những ảnh xám bình thường khác, ta tìm đường biên bằng toán tử sobel.

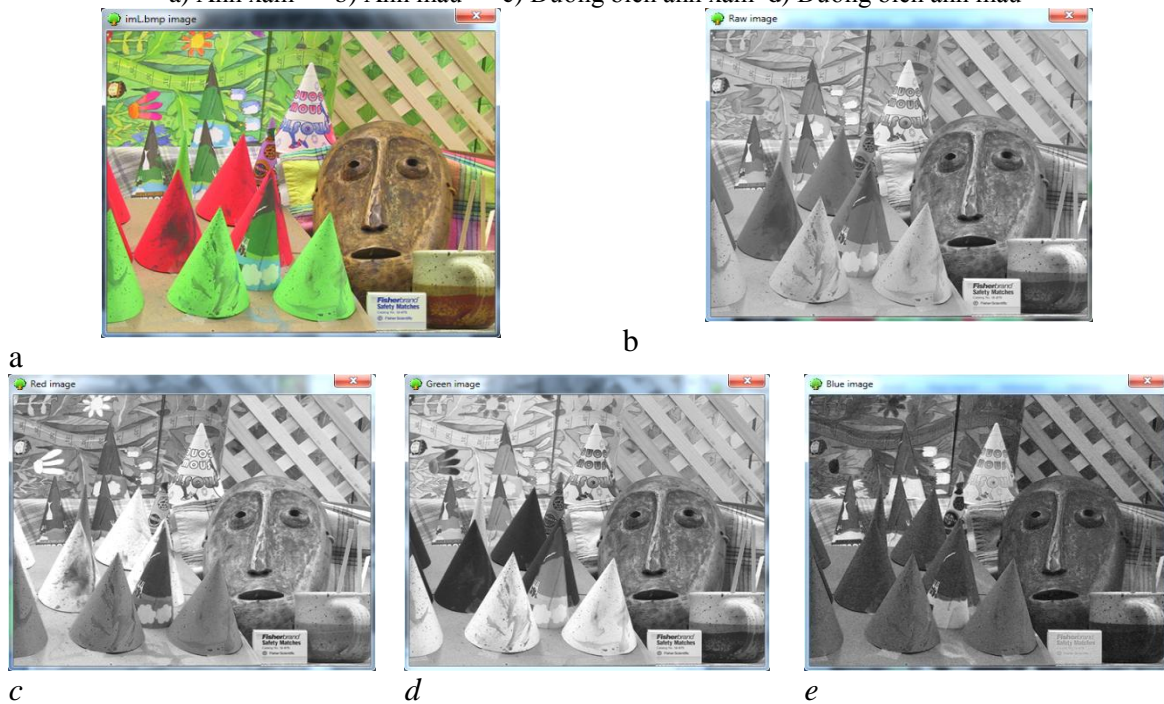
* Bước 2: Ứng với mỗi điểm chỉ có thể có 1 giá trị nên ta sẽ lấy giá trị đường biên của nó bằng giá trị lớn nhất trong 3 giá trị tính được trên 3 mảng màu.

Tính toán SAD dựa trên 3 mảng màu sẽ phải tốn thời gian gấp 3 lần so với tính toán SAD dựa trên ảnh xám. Thật vậy, nếu coi mỗi màu cơ bản tương đương với một ảnh xám, ta sẽ phải tính SAD dựa trên 3 tấm ảnh so với một tấm. Như vậy, dựa theo công thức (3-7), kích thước cửa sổ W sẽ tăng gấp 3 khiến thời gian thực hiện phép toán t cũng tăng gấp 3 lần, tốc độ phép toán sẽ giảm đi 3 lần.

Tuy nhiên, dựa vào phương pháp này, tốc độ phép toán sẽ không bị giảm đi 3 lần như so với phương pháp SAD ảnh xám thông thường. Thật vậy, bởi lẽ mặc dù đường biên được lấy dựa trên 3 mảng màu, mỗi điểm chỉ là đường biên của một mảng màu cơ bản. Như vậy, phép tính SAD dựa vào điểm đó cũng chỉ phải tính trên một mảng màu. Và điều này khiến tốc độ thuật toán lúc này không bị tăng gấp 3 lần.



Hình 17. Đường biên 20% số điểm bằng phương pháp Sobel màu và xám.
 a) Ảnh xám b) Ảnh màu c) Đường biên ảnh xám d) Đường biên ảnh màu

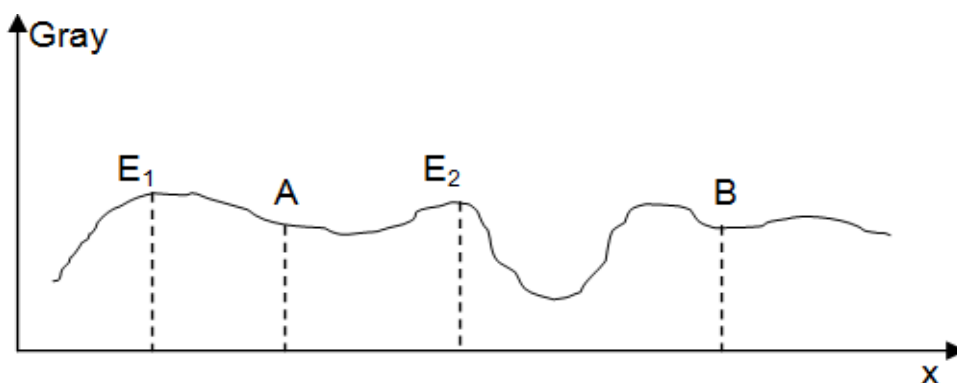


Hình 18. Các mảng màu của hình gốc.
 a) Ảnh màu , b) Ảnh xám , c) Ảnh màu đỏ , d. Ảnh màu xanh lục e. Ảnh màu xanh lam

5. Phương pháp SAD phân vùng

SAD phân vùng là kỹ thuật sử dụng để nội suy một bản đồ độ sâu bao gồm các điểm đã tính được chuẩn. Sử dụng kỹ thuật này trong trường hợp chúng ta có thể đoán được phạm vi của những điểm cần tính. Giả sử một điểm đã được xác định bên trong một vật thể mà những

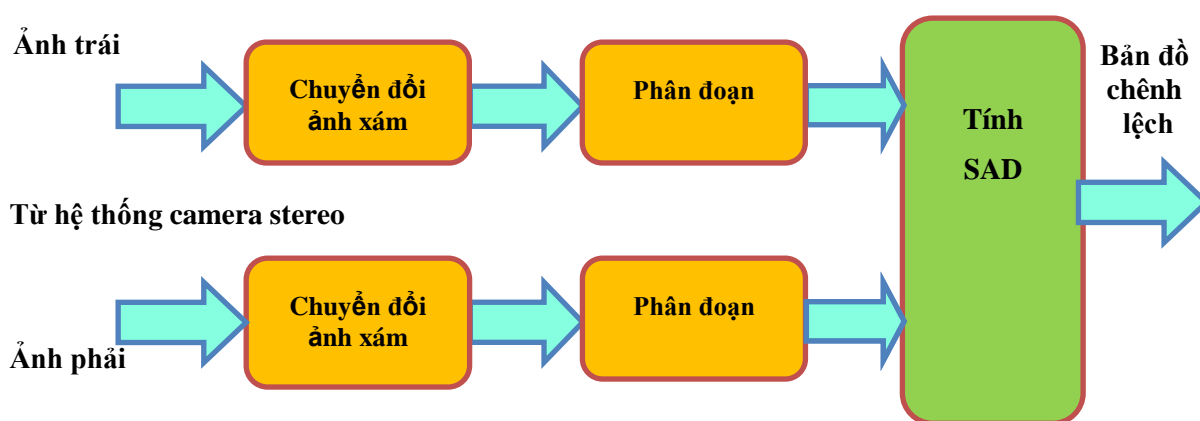
đường biên của nó đã được xác định độ sâu. Khi đó, tồn tại một xác suất cao rằng những điểm bên trong vật thể sẽ có độ sâu trong một phạm vi nhỏ xung quanh giá trị độ sâu của các đường biên của nó. Khi đó, sự giới hạn độ sâu tính toán của các điểm nội suy có thể làm mất khả năng nhầm lẫn giữa điểm đã cho và các điểm bên ngoài.



Hình 19. SAD phân đoạn.

Ở (Hình 19) , giả thiết độ sâu của những điểm thuộc đường biên bao gồm E1 và E2 đã tính được và tính chuẩn. Khi đó, giá trị về độ sâu của A chỉ có thể dao động xung quanh các

giá trị độ sâu của E1 và E2 dẫn đến khả năng tính toán độ sâu của A được nâng cao lên, mặt khác lại không bị nhầm lẫn với một điểm cũng có cấu trúc màu tương tự là điểm B.



Hình 20. Sơ đồ khối tổng quát của phương pháp SAD phân vùng.

Giả thiết biện pháp này không được sử dụng, khi đó ở ảnh đối chiếu có thể xuất hiện một điểm giống với điểm cần tính nằm ở lân cận khoảng cách với điểm B. Khi này, thuật toán SAD có thể đưa đến sự nhầm lẫn ở giá trị độ sâu của A. Điểm A có thể được tính nhầm và ảnh hưởng trực tiếp đến các điểm nằm bên cạnh nó.

Tóm lại, phương pháp SAD phân đoạn dùng tốt trong trường hợp nội suy dựa theo giá trị đã tính đúng. Sử dụng biện pháp này giúp giảm lỗi và tăng tốc độ. Tính toán bản đồ chênh lệch là một trong những vấn đề quan trọng trong thị giác máy tính 3D. Một trong những phương pháp tương đối mới là Cải thiện tính toán Bản đồ độ sâu từ hình ảnh stereo theo phương pháp lai.

Có nhiều phương pháp tính Bản đồ chênh lệch, ở đây ta nghiên cứu hai phương pháp cải thiện tốc độ tính toán (SAD điều chỉnh và SAD phân vùng).

6. Belief Propagation

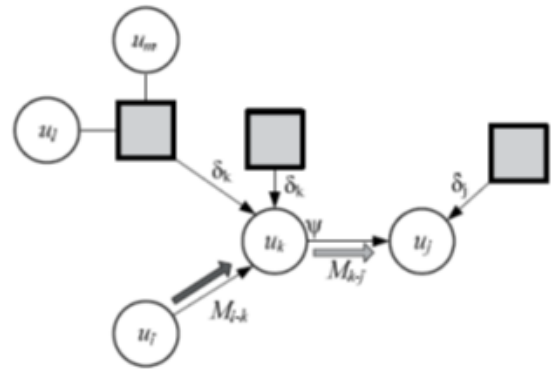
Belief Propagation là các thuật toán đã được phát triển chủ yếu là với mục đích để tìm xác suất cận biên trong mạng Bayes. Ngoài ra, các thuật toán cũng có thể xử lý các mô hình đồ họa khác như mô hình Markov Random Field (MRF), đây là mối quan tâm nhất định trong việc tối ưu hóa các hàm năng lượng toàn cục được tìm thấy trong máy tính thị giác. Mô hình MRF là mô hình đồ thị vô hướng, trong đó các nút đại diện cho các biến ngẫu nhiên. Xác suất chung $P(x_1, \dots, x_n)$ của mô hình MRF pair-wise có thể được viết phân tích như sau:

$$P(x_1, \dots, x_n) = \frac{1}{Z} \prod_i \phi_i(x_i, y_i) \prod_{i,j} \psi_{ij}(x_i, x_j) \quad (6-6)$$

Ở đây x_n đại diện cho các nút của đồ thị, Z là hằng số chuẩn hóa và các sản phẩm trên ij là

hơn các nước láng giềng gần nhất trên mạng vuông. Tiềm năng $\phi_i(x_i, y_i)$, biểu diễn cho xác suất cho một trạng thái nhất định $x_i \in X_i$ trong nút i dựa trên quan sát y_i và tiềm năng $\psi_{ij}(x_i, x_j)$ biểu thị sự phụ thuộc điều kiện giữa các nút lân cận. Biến quan sát y_i hiếm khi được viết ra một cách rõ ràng [15].

Thuật toán Belief Propagation (BP) thể hiện trong (Hình 21) các thông điệp được chuyển đi trong một mô hình đồ họa thông qua một loạt các bản tin được gửi đi giữa các nút lân cận xung quanh trong một lưới hình ảnh liên kề 4. Bản tin được cập nhật trong các lần lặp. Trong mỗi bước lặp, mỗi điểm ảnh của đồ thị liên kề tính toán thông điệp của mình, dựa trên kết quả của bước lặp trước, và gửi tin nhắn mới của nó tới tất cả các điểm ảnh liên kề 4 (hàng xóm) theo hình thức song song [15].



Hình 21. Mô hình cơ bản cho phân vùng ảnh.

Mô hình cơ bản cho phân vùng ảnh, như thể hiện trong (Hình), là một đồ thị với hai loại nút: nút ẩn (vòng tròn) và các nút quan sát (hình vuông). Ψ là hàm chuyển trạng thái giữa một cặp nút trạng thái ẩn khác nhau và δ là hàm đo lường giữa các nút trạng thái ẩn và nút dữ liệu quan sát được. Trong BP, mỗi nút sẽ gửi một thông điệp tới các nút láng giềng trong đó biểu diễn phân bố xác suất [15].

7. Phương pháp Mean Shift:

Thuật toán Mean Shift [16] đã được đề xuất bởi Fukunaga và Hostetler. Các thuật toán dựa trên mật độ ước lượng hạt nhân [17]. Thuật toán này là thuật toán lặp phi tham số. Cho một tập hợp các điểm dữ liệu d-chiều được đại diện bởi các giá trị $x_i, i = 1, 2, \dots, n$ trong d-chiều không gian R^d . Số lượng của điểm x_i thuộc khu vực d-chiều xung quanh x với cạnh dài h được cho bởi:

$$\sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \tag{7-7}$$

trong đó h xác định bán kính của hạt nhân và $K(x)$ là hạt nhân hoặc hàm cửa sổ. Như vậy, ước tính mật độ hạt nhân được cho bởi:

$$f_{h,K} = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \tag{7-8}$$

Trong đó $K(x) = c_{k,d} \cdot k(\|x\|_2)$ là hạt nhân đối xứng xuyên tâm, $k(x)$ là hồ sơ hạt nhân và $c_{k,d}$ được chuẩn hóa liên tục. Sau đó, phương trình (7-10) có thể được viết lại để:

$$\nabla f_{h,K}(x) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x - x_i}{h}\right\|_2\right) \tag{7-9}$$

Ước tính này có thể được xem như là gradient của xấp xỉ mật độ:

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x - x_i) k'\left(\left\|\frac{x - x_i}{h}\right\|_2\right) \tag{7-10}$$

Một hạt nhân mới được định nghĩa là $G(x) = c_{k,d} \cdot g(\|x\|_2)$, trong đó $g(x) = -k'(x)$ là hàm hạt nhân mới và $c_{k,d}$ là hằng số được chuẩn hóa. Sau khi thay thế trong phương trình (7-12):

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x - x_i) g\left(\left\|\frac{x - x_i}{h}\right\|_2\right) \tag{7-13}$$

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n g_i \left(\frac{\sum_{i=1}^n x_i g_i}{\sum_{i=1}^n g_i} - x \right)$$

$$g_i = g\left(\left\|\frac{x - x_i}{h}\right\|_2\right)$$

Sử dụng phương trình (7-13) thay đổi trung bình được xác định:

$$m_{h,g}(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x - x_i}{h}\right\|_2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x - x_i}{h}\right\|_2\right)} - x \tag{7-14}$$

Các thuật toán Mean Shift dựa trên việc tính toán lặp đi lặp lại vector dịch trung bình và thay đổi có ý nghĩa và liên tục hiện thực hóa vị trí hạt nhân bằng phương trình [15],[16]. $x_{k+1} = x_k + m(XK)$. (7-15)

8. Thuật toán SAD phân vùng lại:

Trong phân vùng ảnh, phương pháp lai kết hợp hai hoặc nhiều hơn các thuật toán phân vùng ảnh khác nhau [18]. Ở đây chúng ta xem xét các thuật toán lai được tạo ra bởi sự kết hợp của hai kỹ thuật: Belief Propagation và các thuật toán phân đoạn Mean Shift nhanh chóng và rất chính xác.

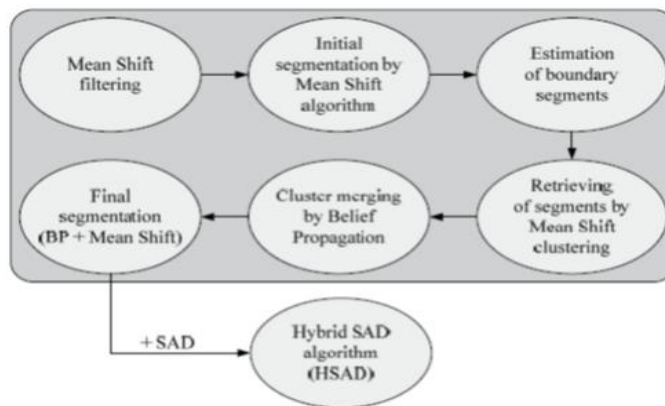
Đầu tiên, chúng tôi áp dụng cách lọc hình ảnh bằng thuật toán Mean Shift. Bước này rất hữu ích cho việc loại bỏ tiếng ồn, làm mịn và phân vùng ảnh. Đối với mỗi điểm ảnh của một hình ảnh, tập hợp các pixel lân cận được xác định. Cho X_i là đầu vào và Y_i hình ảnh lọc, trong đó $i = 1, 2, \dots, n$. Các thuật toán lọc bao gồm các bước sau đây [19],[20],[21]:

* Khởi tạo $j = 1, y_{i,1} = p_i$.

* Tính toán thông qua Shift bình chế độ, nơi hội tụ điểm ảnh.

* Lưu trữ các thành phần của mức xám của các giá trị tính toán $Z_i = (x_i, y_i, c)$ tại Z_i , nơi x_i là thành phần không gian và y_i, c là thành phần phạm vi.

Thứ hai, hình ảnh được chia thành các đoạn



Hình 22. Thuật toán SAD lai.

Tính toán độ lệch: Trong mục này, chúng ta diễn tả quá trình của sự so sánh chênh lệch nơi ảnh được đưa vào là những mảnh (đoạn) lần đầu tiên và sau đó những điểm có cùng điều kiện ở bên trái và bên phải ảnh được tìm thấy.

Đây là một ý tưởng được giải thích cho vùng 3D duy nhất điểm P trên (hình 23). Để đưa ra khoảng cách của vật được quan sát bởi 2 máy quay cùng nhìn theo một hướng nhưng khoảng cách khác nhau được biết (định nghĩa) là đường cơ bản. Sau đó, vật thể xuất hiện ở vị trí tương tự trong cả hai ảnh 3 chiều. Khoảng cách giữa vật ở bên trái và bên phải ảnh được biết (định nghĩa) là khoảng chênh lệch d bởi (8-16). X_L và X_R là tọa độ x của ước

sử dụng thuật toán phím Mean Shift. Trong bước thứ ba, thực hiện sắp xỉ đường bao của các phân đoạn. Thứ tư, các phân đoạn nhỏ được sáp nhập với nhau thành những đoạn tiếp giáp tương tự như hầu hết các phương pháp Belief Propagation. Cuối cùng, chúng ta tích hợp thuật toán phân đoạn lai đề xuất với thuật toán phù hợp stereo SAD. Phương pháp này SAD lai có thể cung cấp kết quả bản đồ độ sâu chính xác cao.

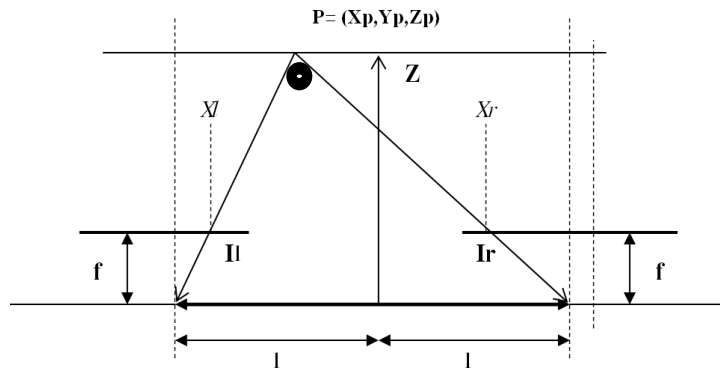
lượng tọa độ 3D còn trên ảnh phẳng bên trái và bên phải là I_L và I_R .

$$d = x_L - x_R = f \left(\frac{x_p + l}{z_p} - \frac{x_p - l}{z_p} \right) \quad (8-16)$$

$$z_p = \frac{2fl}{d} = \frac{fB}{d} \quad (8-17)$$

Khi bên trái và bên phải của ảnh phẳng quay được có cùng vị trí phẳng, tọa độ y của 2 ảnh là bằng nhau ($y_L = y_R$) và có độ chênh lệch tọa độ ngang (trục x) là $(x_L - x_R)$. [22]

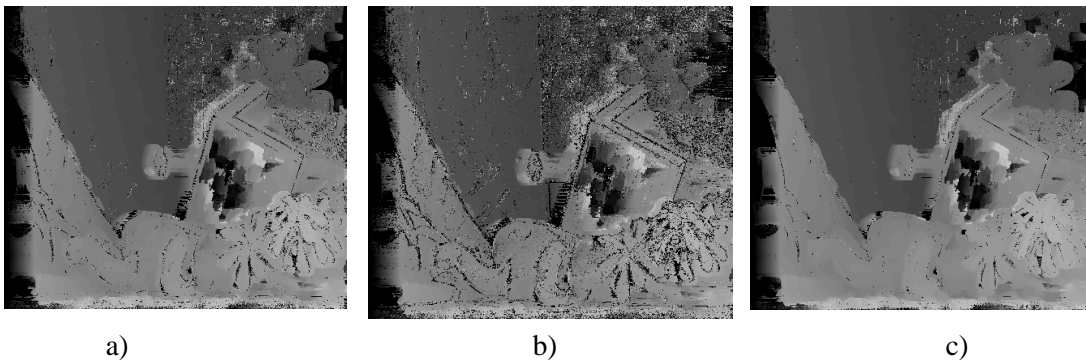
Điều này có nghĩa là độ chênh lệch được tính toán, cường độ có thể thay đổi được khi thay đổi thông số của máy quay: chiều dài tiêu cự f và khoảng cách của đường cơ sở $B = 2l$.



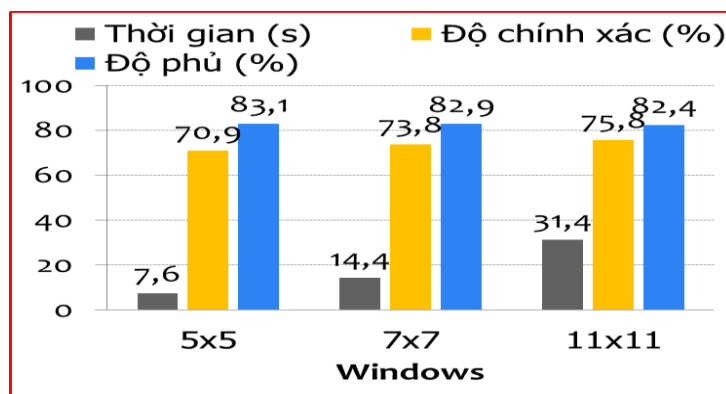
Hình 23. Một hệ thống stereo đơn giản .

9. Nhận xét và đánh giá:

9.1. Phương pháp SAD truyền thống

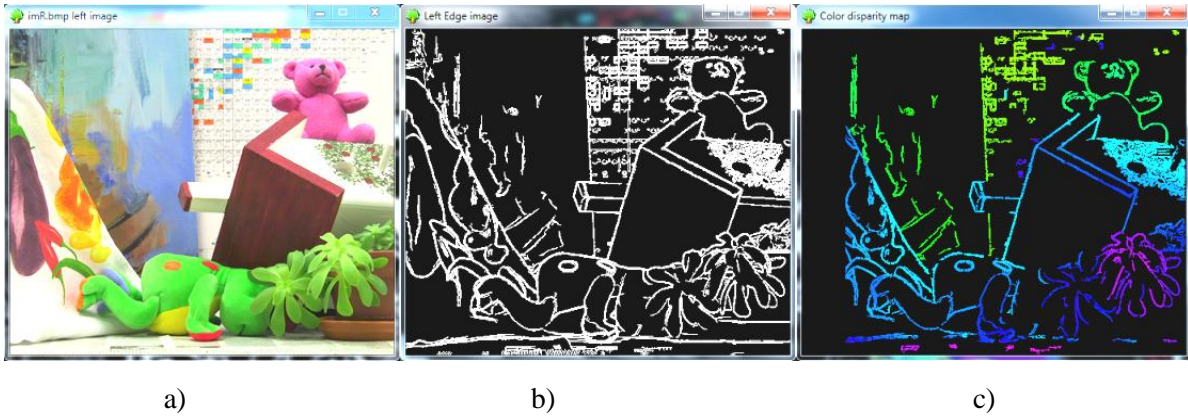


Hình 24. Một số kết quả Bản đồ chênh lệch SAD truyền thống.
 a. Bản đồ chênh lệch với cửa sổ 5x5;
 b. Bản đồ chênh lệch với cửa sổ 7x7; c. Bản đồ chênh lệch với cửa sổ 11x11;

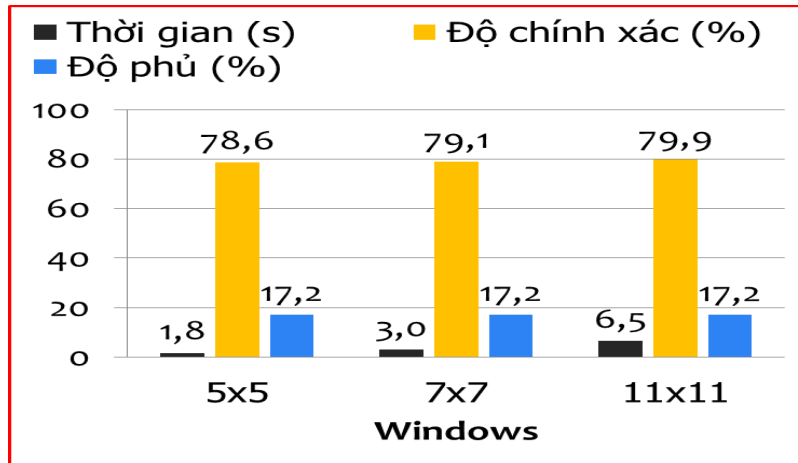


Hình 25. Đồ thị thời gian, độ phù và độ chính xác.

9.2. Phương pháp SAD có điều chỉnh:



Hình 26. Một số kết quả thực hiện tìm SAD trên đường biên.



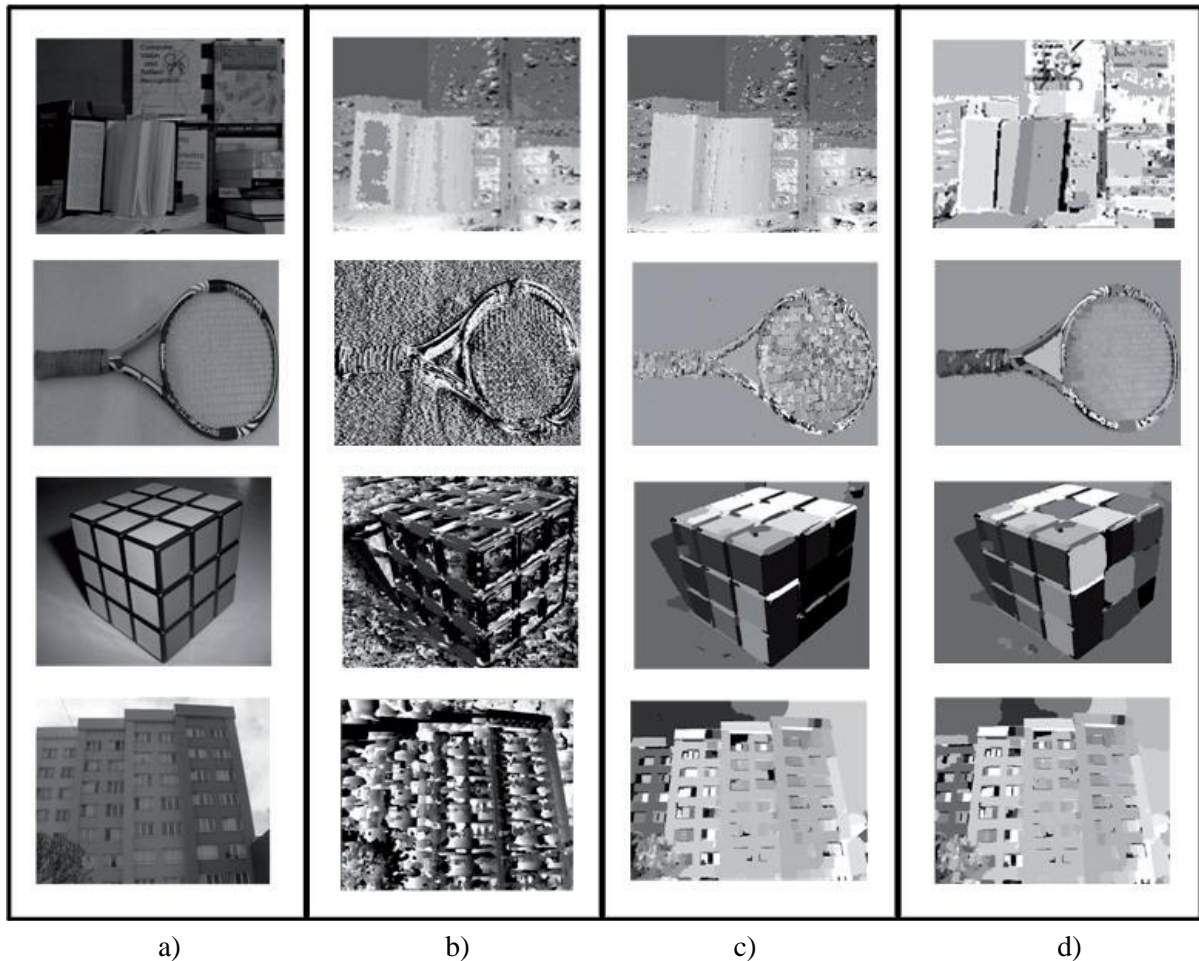
Hình 27. Đồ thị thời gian, độ phủ và độ chính xác.

Thời gian thực hiện tính bản đồ chênh lệch theo phương pháp SAD điều chỉnh đã giảm rất nhiều so với phương pháp SAD thông thường.

c. Phương pháp SAD phân vùng:

Trong cuộc thử nghiệm tiếp theo, các phương pháp lai đề xuất được thử nghiệm trên bốn hình ảnh stereo thực tế được thực hiện bởi hệ thống camera stereo. Thuật toán này được áp dụng để ước tính sự chênh lệch của các khối 9x9 từ hình ảnh bên trái tham chiếu bằng cách

tìm kiếm các khối ứng cử viên tương ứng của hình ảnh bên phải. Bản đồ chênh lệch cuối cùng của bốn hình ảnh lập thể thử nghiệm, cụ thể là sách, vợt, khối lập phương và tòa được thể hiện trong (Hình 28). Cụ thể, các (Hình 28-b) cho thấy các bản đồ khác biệt được tạo ra bởi thuật toán SAD mà không phân đoạn và hình. (Hình 28-c) kết quả theo thuật toán phân chia lại. Bản đồ chênh lệch sản xuất bởi máy ảnh nói Bumblebee được hiển thị trong (Hình 28-d).



Hình 28. Kết quả thực nghiệm trên bốn ảnh thử nghiệm màu xám.

- a) Hình ảnh tham khảo, b) Bản đồ chênh lệch sau khi sử dụng thuật toán SAD mà không phân đoạn
- c) Kết quả của thuật toán SAD lại, d) Hệ thống camera Bumblebee.

Chất lượng của các bản đồ khác biệt được biểu diễn như là tỷ lệ phần trăm của các điểm ảnh với các lỗi sai lệch này (pixel có sự phù hợp xấu) [6]:

$$P = \frac{1}{X * Y} \sum_{i=1}^X \sum_{j=1}^Y (|d_c(i, j) - d_T(i, j)|) \tag{9-18}$$

trong đó $X * Y$ đại diện cho kích thước của hình ảnh, d_c là bản đồ chênh lệch tính của hình ảnh thử nghiệm và d_T là bản đồ chênh lệch sự thật. Sự thật chênh lệch bản đồ mặt đất là

ngược đảo của khoảng cách thật mặt đất. Phương trình (9-19) cho thấy làm thế nào để tính toán bản đồ chênh lệch sự thật mặt đất từ bản đồ độ

$$d_T = \frac{fBI_{RES}}{D_T h} \tag{9-19}$$

với D_T là sự thật bản đồ độ sâu mặt đất, h là chiều cao từ mặt phẳng đất, $D_T * h$ là khoảng cách thật mặt đất, B là cơ sở giữa các máy ảnh, I_{RES} có độ phân giải hình ảnh và f là tiêu cự.

Bảng 1. So sánh kết quả của các giải thuật tính bản đồ chênh lệch

Ảnh kiểm tra	SAD thông thường	SAD phân vùng lai	Hệ thống camera Bumblelee
Quyển sách	9,64%	4,12%	3,59%
Cái vợt	8,48%	4,72%	3,83%
Rubic	9,26%	4,37%	3,51%
Tòa nhà	9,91%	4,21%	3,65%

Các bản đồ chênh lệch thu được từ việc đề xuất lai phân đoạn dựa trên thuật toán SAD phân vùng lai được so sánh với các phương pháp khác (sử dụng hệ thống camera Bumblebee và phương pháp SAD mà không phân đoạn). Trong (bảng 1), sự khác biệt giữa tỷ lệ phần trăm điểm ảnh lỗi chênh lệch được trình bày. Như có thể thấy trong (hình 28), các thuật toán SAD phân vùng lai tạo ra kết quả chính xác hơn so với thuật toán SAD mà không phân đoạn. Ngoài ra, hiệu suất của phương pháp là gần đạt đến hiệu suất của hệ thống camera Bumblebee.

Cuối cùng, các thuật toán phân đoạn lai đã tận dụng được lợi thế nhanh chóng của thuật toán Mean Shift và chính xác từ thuật toán Belief Propagation. Lợi thế lớn của một thuật toán SAD lai là bất kỳ phân đoạn đều được phát hiện được nhận dạng tự động. Thuật toán này đã đạt được một hiệu suất phân đoạn gần thời gian thực và kết quả chính xác cao.

10. Kết luận:

Bài viết nghiên cứu so sánh cho việc khôi phục bản đồ độ sâu dựa trên các thuật toán phân vùng lai được đề xuất sử dụng quá trình lọc. Các thuật toán đã được thử nghiệm trên thực tế với cảnh phức tạp và tình trạng ánh sáng kém dựa trên hệ thống camera stereo công nghiệp từ

Bumblebee, từ kết quả thí nghiệm ta có thể tìm một bản đồ chênh lệch, mà dịch chuyển chênh lệch giữa hai hình ảnh và cuối cùng là được sử dụng để ước tính giá trị chiều sâu. Dựa trên các thuật toán tổng bình phương khác biệt (SSD) và tổng sự khác biệt tuyệt đối (SAD) có thể tìm thấy bản đồ độ sâu trực tiếp. Tùy thuộc vào các kỹ thuật phù hợp dựa trên các thuật toán SAD và SSD một nguyên tắc giữa thời gian chạy và chất lượng cần phải được lựa chọn. Cuối cùng, thuật toán SAD phù hợp hình ảnh stereo được lựa chọn để thực hiện các thí nghiệm tiếp theo, bởi vì các thuật toán SSD đòi hỏi gần như tăng gấp đôi thời gian. Kết quả thí nghiệm được trình bày trong (Hình 28) cho thấy thuật toán SAD phân vùng lai cho kết quả tốt hơn hơn thuật toán SAD mà không cần sử dụng pjaan vùng và SAD điều chỉnh nhưng hơi kém hơn so với máy ảnh nổi của hệ thống Bumblebee. Các thuật toán SAD phân vùng lai được đề xuất cũng cho thấy hiệu suất cao hơn. Trong tương lai, tôi dự định thực hiện nghiên cứu và cũng kiểm tra các thuật toán phức tạp hơn.

Tài liệu tham khảo:

- [1] Dorit Borrmann, Jan Elseberg, Kai Lingemann, Andreas Nüchter (2011), “ The 3D Hough transform for plane detection in point clouds: A Review and a new Accumulator Design”, Received: 13 January 2011 / Revised: 13 February 2011 / Accepted: 10 March 2011, © 3D Research Center and Springer 2011
- [2] Yedida, J. S, Freeman, W. T, Weiss, Y. Understanding belief propagation and it is generalizations. Exploring Artificial Intelligence in the New Millennium, 2003, Chap.8, p, 236 – 239.
- [3] POINT GREY, Bumblebee Stereo Vision Camera Sytems. BB2 – 08S2 (datasheet). 2 pages. [Online] Cited 2011-10-23. Available at: http://www.ptgrey.com/products/bumblebee2_xb3_datasheet.pdf
- [4] Scharstein, D, Szeliski, R, A taxonomy and evaluation of dense two frame stereo

- correspondence algorithms. *International Journal of Computer Vision*, 2003, vol, 47, no. (1/2/3), p.& - 42
- [5] Brown, M, Z, Burschka, D, Hager, G, D, Bhartia, P. Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, vol, 25, no.8.
- [6] KUHL, A. Comparison of stereo matching algorithms for mobile robots. *Centre for Intelligent Information Processing Sytem 2005*, University of Western Australia, p. 4 – 24.
- [7] Cao, X., Foroosh, H, H. Camera calibration using symmetric objects. *Image Processing IEEE Transactions*, 2006, vol, 15, no. 11, p.3614 – 3619.
- [8] Shimizu, M., Okutomi, M, Calibration and rectification for reflection stereo. In *IEEE Conference on Computer Vision and Pattern Recognition CVPR Anchorage (USA)*, 2008, p.1 – 8.
- [9] Hartley, R, Zisserman, A, *Multiple View Geometry in Computer Vision 2nd ed.* Cambridge: University Press, 2005.
- [10] Zhang, Y. J. *Advance in Image and Video Segmentation*. 1nd ed. USA: IRM Press, 2006.
- [11] Wang, G, JU, H, A disparity map extraction algorithm for lunar rover BH 2. In *IEEE International Conference on Intelligent Computing and Intelligent Sytems ICIS 2009*. Shanghai, 2009, vol. 4, p. 385 – 389.
- [12] Mckinnon, B., Baltas, J. Practical region – based matching for stereo vision. In *IWCIA, 2005*, vol. 3322 of *Lecture Notes in Computer Science*, Springer, p. 726 – 738.
- [13] Craig Watman, “Fast Sum of Absolute Differences Visual Landmark Detector”, *Robotic Systems Laboratory Department of Systems Engineering, RSISE, Australian National University, Canberra, ACT 0200 Australia*, Email: d.austin@computer.org
- [14] Christor Georgoulas, Georgios CH, Real-time stereo vision applications. *Laboratory of Electronics, Democritris University of Thrace Xanthi, Greece* p.267-292.
- [15] Guan, S, Klette, R, Belief Propagation on edge image sequences. In *Proceedingx Robot Vision LNCS 4931*, 2006, p. 291 – 302.
- [16] Siqiang, L., Wei, L, Image segmentation based on the Mean – Shift in the HSV space. In *26th Chinese Control Conference*, 2007. [Online], p. 476 – 479.
- [17] Benco, M., Hudec, R, The advances image segmentation techniques for broadly useful retrieval in large image database. In *NSSS IX, Tatranske Zruby (Slovakia)*, 2006, p. 40 – 44.
- [18] He, R, Zhu, Y. A hybrid image segmentation approach based on Mean Shift and fuzzy C – Means. In *Asia Pacific Conference on Information Processing*. [Online], 2009.
- [19] Comaniciu, D., Meer, P. Mean Shift: A robust approach toward feature space analysis *IEEE Transaction on Pattern Analysis and Machine Intel.*, 2002, vol. 24, no. 5, p. 1 – 18.
- [20] Rodriguez, R., Suarez, A. G. An image segmentation algorithm using iteratively the Mean Shift. *Progress in Pattern Recognition, Image Analysis and Applications*. 2006, LNCS 4225, Berlin Heidelberg: Springer, vol. 4225/2006, p. 326 – 335.
- [21] Hudec, R. *Adaptive Order – statistics L – filters*, 1st ed. University of Zilina: EDIS Press, 155 pages, ISBN 978 – 80 – 554 – 0248 – 2, 2011 (in Slovak).
- [22] Rostam Affendi Hamzah, Hasrul Nisham Rosly, “An Obstacle Detection and Avoidance of A Mobile Robot with Stereo Vision Camera” , *Faculty of Electronics & Computer Engineering, UTeM, Malaysia*. rostamaffendi@utem.edu.my

Comparative Study of Image Processing Algorithms Calculate Depth Stereo Image Camera Applications in the Visual System

Bạch Ngọc Minh

Centre Development Vietnam National University, Hanoi, 144 Xuân Thủy, Hanoi, Vietnam

Abstract: There are three basic methods to compare the similarity between two images, which is SAD, SSD and NCC. NCC method is using cross- correlation between the pixels so the calculations are complex . These include the arithmetic average, total, multiplication total, square root should this method definitely perform very time consuming. Both SAD method and SSD method does'nt much difference the complexity of the algorithm so we need to do an experiment to verify the quality of the depth maps and the speed of execution. From the results that we choose a more appropriate method

Comparison of running time between SAD and SSD algorithms show that the choice of the SAD algorithm is more reasonable, because the SSD algorithm needs nearly twice the time to run. SAD method can solve the problem of finding depth in 3D images. A point in space to acquire images of two lenses in two different locations. Therefore so before finding the depth of the pixel, we need to find the corresponding pixels. SAD method is the easiest method to find the pixel .

SAD method has many calculate ways, calculate by traditional methods SAD, adjustment SAD method (based on the image border) , segmentation SAD method (image segmentation). In Here we choose between two methods of adjustment SAD method and segment SAD method decreased 20 % compared with traditional SAD.

Then we consider the hybrid algorithm SAD partition is created by the combination to two techniques: Belief Propagation and the Mean Shift segmentation algorithm. This approach combines the advantages of both methods segment. Mean Shift algorithm is fast results and Belief Propagation perform very accurate segmentation.

Keywords: Image segmentation, Belief Propagation, Mean Shift, SAD, 3D image.