NGHIÊN CỨU KẾT HỢP KỸ THUẬT THỊ GIÁC (VISION) VÀ TRUYỀN NHẬN KHÔNG DÂY TRONG NHIỆM VỤ DẪN ĐƯỜNG CHO ROBOT

Vương Ngọc Dũng
vndung@dme.hcmut.edu.vn, v_n_dung@yahoo.com
Khoa Cơ khí, Trường Đại học Bách Khoa TP.HCM

ABSTRACT

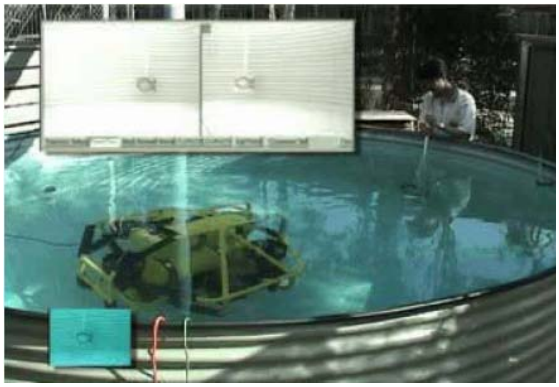
This paper mentions to a low-cost stereo system for mobile robot: design, build and investigate. It will also introduce to some basic theories for processing the information from this system.

TÓM TẮT

Bài báo này sẽ đề cập đến việc xây dựng và phát triển một hệ thống stereo di động (nhằm sử dụng cho robot) có giá thành thấp. Đồng thời cũng sẽ trình bày tóm tắt lý thuyết liên quan đến các kỹ thuật xử lý các thông tin nhận được từ hệ camera, hiện thực để khảo sát mô hình xây dựng.

I. TỔNG QUAN

Cho đến thời điểm hiện tại, đa phần các robot khi hoạt động đều cần có sự giúp đỡ của con người dưới một hình thức nào đó. Vì vậy, việc hỗ trợ cho con người trong quá trình điều khiển là một trong những vấn đề đang được quan tâm hiện nay. Lấy ví dụ robot mà ta đang quan tâm là một robot hoạt động dưới nước. Do đặc điểm của môi trường nước nên việc tự động hoàn toàn các hoạt động của robot khi thám hiểm là một vấn đề nhiều rủi ro. Vì vậy đa phần các robot thám hiểm dưới nước đều được điều khiển từ xa. Robot được trang bị camera để người điều khiển có thể có tầm quan sát rộng hơn. Ngoài việc mở rộng tầm nhìn, ta còn có thể sử dụng hệ camera này để đưa ra các thông tin hữu ích hơn cho người điều khiển.



H1.1 Robot hoạt động trong môi trường nước.

Ở đây, hệ camera được sử dụng song song với các loại cảm biến khác (như cảm biến siêu âm để đo khoảng cách, cảm biến va chạm, cảm biến la bàn,...) chứ không thể thay thế hoàn toàn các loại cảm biến khác vì các ưu khuyết điểm của chúng sẽ bổ xung lẫn nhau. Đây là xu hướng phát triển tất yếu của các robot trong tương lai – không phủ định (bỏ đi) các loại cảm biến truyền thống cũ mà kết hợp chúng với nhau (dùng song song) – "sensor fusion"

Tuy nhiên, chi phí cho việc trang bị hệ thống camera nêu trên thường rất cao (khoảng trên 10.000 USD cho các hệ thống chuyên nghiệp) nên việc nghiên cứu sẽ gặp nhiều khó khăn, nhất là trong điều kiện kinh tế hiện nay. Xuất phát từ thực tế trên, bài báo sẽ đề cập đến việc xây dựng và phát triển một hệ thống stereo (hệ chỉ dùng hai camera) cho robot di động hoạt động **trong nhà** với các mục tiêu đặt ra như sau:

- Có tính **di động** cao.
- Giá thành **thấp**.

II. XÂY DỰNG MÔ HÌNH

Một hệ thống stereo cơ bản có thể sử dụng để hỗ trợ việc điều khiển robot di động tối thiểu phải có các thành phần sau:

- Hai camera chuyên dùng.
- Hệ thống cơ giúp thay đổi và mở rộng thị trường quan sát của hệ camera.
- Hệ thống điều khiển vị trí hệ thống cơ.
- Hệ thống truyền nhận không dây (vì đối tượng phục vụ của hệ thống là robot di động) dữ liệu và thông tin điều khiển của hệ thống stereo.

Và giá thành của một hệ thống như vậy hiện nay trên thế giới vào khoảng từ 10.000 USD cho đến hàng trăm ngàn USD [4]. Điều này sẽ gây khó khăn không ít cho việc tìm hiểu và thực nghiệm các lý thuyết điều khiển trên các robot vision. Như mục tiêu ban đầu của đề tài, tác giả đã tìm hiểu và xây dựng một hệ thống stereo có giá thành thấp nhưng vẫn đáp ứng về cơ bản các yêu cầu nêu trên như sau:

- Sử dụng một thiết bị thông dụng là máy tính đóng vai trò giao tiếp trung tâm của hệ stereo.

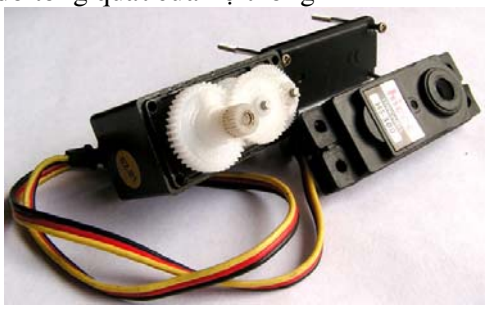
- Sử dụng các camera dân dụng/ webcam kết hợp với bộ phần mềm bắt hình ở mức hệ thống. Ở đây ta có thể kết hợp với các bộ thu và giải mã hình ảnh bằng phần cứng để tăng tốc độ thu nhận. Với một máy tính và bộ thu nêu trên, tốc độ bắt hình hoàn toàn có thể đạt đến 30 fps.

- Sử dụng mô hình mạng không dây dân dụng làm hệ thống trao đổi thông tin vô tuyến. Điều này là hợp lý vì môi trường làm việc của các robot di động là trong nhà hoặc công sở như đã đề cập từ đầu.

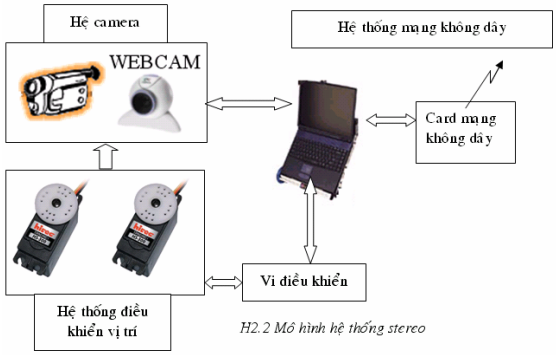
- Do yêu cầu của hệ thống cơ là điều khiển chính xác vị trí của hệ camera nên ở đây ta có thể sử dụng động cơ RC Servo để điều khiển vị trí thông qua độ rộng xung cấp vào (PWM).

- Hệ thống cơ được cấu tạo từ các động cơ RC Servo được điều khiển bằng một vi điều khiển và vi điều khiển này sẽ nhận tín hiệu điều khiển từ máy tính để thay đổi vị trí của hệ camera.

Sơ đồ tổng quát của hệ thống



H2.1 Động cơ RC Servo



H2.2 Mô hình hệ thống stereo

Từ ý tưởng nêu trên, ta hiện thực được mô hình như sau:

- 2 Creative Webcam Pro PD1030DV giao tiếp USB 1.1
- 2 động cơ RC servo Hitec HS-300 được điều khiển bằng một vi điều khiển PIC giao tiếp nối tiếp với máy tính.
- LinkSYS Wireless card, chuẩn 802.11g, tần số 2.4 GHz
- Acer laptop PII, 300MHz, 256M.

Với giá thành vào khoảng 500 USD.

H2.2 Mô hình hệ thống stereo thực tế

Việc xây dựng hệ thống stereo như trên còn có lợi điểm là do robot được trang bị laptop nên ta có thể dùng một card mạng không dây PCMCIA chuẩn 802.11g. Đây là một trong những chuẩn không dây dân dụng mới nhất. Tốc độ truyền có thể lên tới 108Mbps và khoảng cách có thể đạt tới 300 mét tùy công suất phát. Lưu ý là tốc độ truyền sẽ tỉ lệ nghịch với khoảng cách giữa hai thiết bị:

Khoảng cách truyền	Tốc độ truyền
60m	54Mbps
70m	48Mbps
80m	36Mbps
120m	24Mbps – 6Mbps
200m	2Mbps
300m	1Mbps

H2.3 Thông số của card mạng không dây

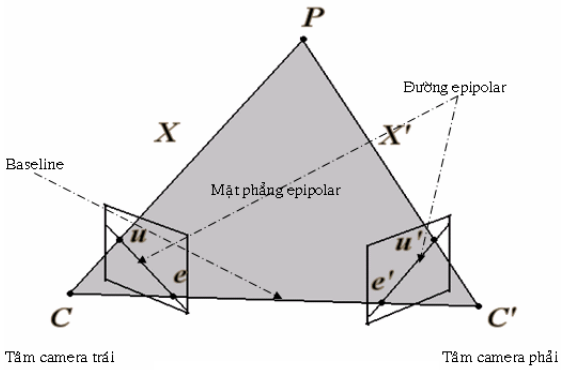
Vì vậy phương pháp truyền nhận không dây này chỉ thích hợp dùng trong các **văn phòng, cao ốc được xây dựng sau này (thường đã được trang bị các điểm tiếp sóng trong tòa nhà)**.

Một thuận lợi khác của mô hình trên là ta có thể dùng bất cứ một máy tính nào có hỗ trợ truyền nhận mạng không dây (hoặc đơn giản chỉ là một máy tính có trang bị card mạng không dây) cùng chuẩn 802.11g và đang kết nối. Điều này làm **tăng tính linh hoạt** của hệ thống nêu trên lên rất nhiều vì ta đang tận dụng hệ thống mạng cục bộ của nơi làm việc nên nếu hệ thống mạng này có kết nối Internet thì ta có thể dễ dàng điều khiển robot **từ bất cứ nơi nào** trên thế giới. Và đây cũng là một trong những điểm mà bài báo muốn nhấn mạnh: "*Nghiên cứu kết hợp kỹ thuật thị giác (vision) và truyền nhận không dây (wireless) trong nhiệm vụ dẫn đường cho robot*".

III. HỆ THỐNG STEREO

Phần này sẽ trình bày sơ lược các khái niệm cơ bản của một hệ thống stereo trong trường hợp tổng quát. Thông tin chi tiết của quá trình tính toán và kiểm tra các thông số cho hệ thống stereo xin tham khảo tài liệu [1][2].

1. Mô hình



H3.1 Mô hình stereo

Mô hình đơn giản nhất là hai camera giống nhau được đặt cách nhau một khoảng (baseline distance) là b. Các mặt phẳng ảnh thì đồng phẳng với nhau trong mô hình này. Một đối tượng trong không gian được quan sát bởi hai camera ở những vị trí khác nhau trong mặt phẳng ảnh. Độ lệch giữa các vị trí của hai đối tượng trong mặt phẳng ảnh được

gọi là disparity. Mặt phẳng đi qua hai tâm camera và điểm vật trong không gian gọi là epipolar plane. Giao tuyến của epipolar plane với mặt phẳng ảnh là epipolar line.

Ta định nghĩa hai khái niệm sau:

- Một cặp điểm ảnh liên hợp (a conjugate pair): là hai điểm ảnh của cùng một điểm vật trong hai ảnh khác nhau.
- Độ lệch (disparity): là khoảng cách giữa hai điểm của một cặp điểm liên hợp khi hai ảnh này xếp chồng lên nhau.

Trong hình vẽ 3.1 trên điểm P có hai ảnh là u và u' nằm trên hai mặt phẳng ảnh. Không mất tính tổng quát ta giả sử rằng tâm của gốc tọa độ trùng với tâm của thấu kính trái.

Gọi z là tọa độ của điểm vật P theo phương z, b là khoảng cách giữa hai camera, f là tiêu cự chung của cả hai camera, u và u' là ảnh của P trên hai camera trái và phải. Ta có thể dễ dàng xây dựng được công thức tính khoảng cách z trong trường hợp hai trục quang của hai camera song song và vuông góc với đường baseline bằng các định lý hình học như sau:

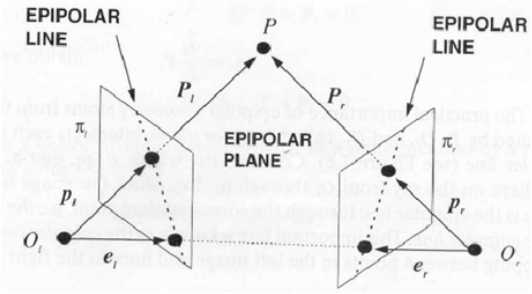
$$z = \frac{bf}{(u - u')}$$

Vì vậy khoảng cách ở những điểm khác nhau có thể được phát hiện thông qua disparities của những điểm ảnh tương ứng.

Chú ý rằng do tính chất rời rạc của ảnh số, các giá trị disparity là các số nguyên ngoại trừ các thuật toán đặc biệt được sử dụng để tính toán với độ chính xác rất cao. Vì vậy với một camera đã biết các thông số, độ chính xác của việc tính toán khoảng cách đối với một điểm vật sẽ tăng khi ta tăng độ dài baseline b. Khi đó độ lệch disparity của một cặp ảnh liên hợp lớn.

2. Các mối quan hệ cơ bản của hệ thống stereo

Xét mô hình hai camera được biểu diễn như sau:



H3.2 Mô hình stereo camera

Hai camera có mối quan hệ với nhau bởi vectơ tịnh tiến $T = (O_r - O_1)$ và một ma trận quay R.

Gọi P_1 và P_r lần lượt là tọa độ của P trong hai hệ tọa độ trái và phải. Thế thì giữa chúng có mối quan hệ như sau:

$$P_r = R(P_1 - T)$$

Gọi p_1 và p_r là tọa độ của P trên hai mặt phẳng ảnh, ta có mối liên hệ giữa tọa độ thực và tọa độ ảnh như sau:

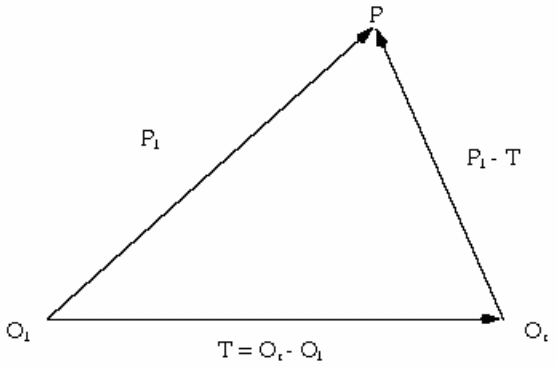
$$p_1 = \frac{f_1}{Z_1} P_1$$

$$p_r = \frac{f_r}{Z_r} P_r$$

MA TRẬN E

Giả sử rằng hệ tọa độ thực trùng với hệ tọa độ của camera trái. Từ điều kiện đồng phẳng của mặt phẳng epipolar plane ta có:

$$(P_1 - T)^T (T \times P_1) = 0 \quad (3.1)$$



H3.3 Mô hình đơn giản của hai camera

Do $P_r = R(P_1 - T)$ và $R^T = R^{-1}$ nên

$$(R^T P_r)^T (T \times P_1) = 0$$

Biểu diễn tích hữu hướng $T \times P_1$ dưới dạng ma trận ta có:

$$T \times P_1 = SP_1$$

trong đó:

$$S = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$

Như vậy ta có:

$$(R^T P_r)^T SP_1 = 0$$

$$P_r{}^T RSP_1 = 0$$

Ma trận S luôn luôn là ma trận không đầy đủ và nó có hạng là rank(S) =2.

Từ: $p_1 = \frac{f_1}{Z_1} P_1$

$$p_r = \frac{f_r}{Z_r} P_r$$

$\Rightarrow$ $P_1 = \frac{Z_1}{f_1} p_1$

$$P_r = \frac{Z_r}{f_r} p_r$$

Thay P_1 và P_r vào công thức: $P_r^T RSP_1 = 0$ ta được:

$$\frac{Z_r}{f_r} p_r{}^T RSp_1 \frac{Z_1}{f_1} = 0$$

$$p_r{}^T RSp_1 = 0$$

Đặt $E = RS$: **E được gọi là essential matrix**

Cuối cùng ta được:

$$p_r{}^T Ep_1 = 0 \qquad (3\text{-}5)$$

Ma trận E có các tính chất là chỉ chứa thông tin của các thông số ngoại của hệ camera. (3-6)

MA TRẬN F

Gọi M_1 và M_r lần lượt là các ma trận thông số nội của camera trái và camera phải, p_1' và p_r' lần lượt là các tọa độ pixel trên ảnh trái và ảnh phải. Thế thì ta có mối quan hệ sau:

$$\begin{aligned} p_1' &= M_1 p_1 \\ p_r' &= M_r p_r \\ p_1 &= M_1{}^{-1} p_1' \\ p_r &= M_r{}^{-1} p_r' \end{aligned} \qquad (3\text{-}7)$$

Thay vào phương trình $p_r{}^T Ep_1 = 0$ ta được:

$$(M_r{}^{-1} p_r')^T E(M_1{}^{-1} p_1') = 0$$

$$p_r{}'^T (M_r{}^{-1})^T EM_1{}^{-1} p_1' = 0 \; p_r{}'^T$$

Đặt:

$$F = (M_r{}^{-1})^T EM_1{}^{-1} = (M_r{}^{-1})^T RSM_1{}^{-1} \qquad (3\text{-}8)$$

F được gọi là fundamental matrix

thì ta nhận được:

$$p_r{}'^T Fp_1' = 0$$

Ma trận F có các tính chất là mang cả thông tin của thông số ngoại và thông số nội của hệ camera.

Có nhiều giải thuật để tìm ma trận F được đưa ra trong những năm gần đây. Cơ bản chúng đều dựa trên ý tưởng của thuật toán 8 điểm.

Để kiểm tra khả năng sử dụng của mô hình đề ra, tác giả đã xây dựng một bộ phần mềm để thu nhận và tính toán các thông số cơ bản của hệ thống stereo đang xây dựng. Quá trình tính toán và giải thuật chi tiết xin tham khảo tài liệu [2].

III. THỰC NGHIỆM

Từ hệ thống stereo được xây dựng như trên. Sử dụng các lý thuyết về kỹ thuật thị giác và kỹ thuật truyền thông trên mạng cục bộ, các kết quả thực nghiệm các thông số kỹ thuật của mô hình như sau (thông tin về quá trình hiệu chuẩn xin tham khảo tài liệu [1]):

Ma trận thông số nội của camera trái và phải:

```
M = [389.152771 , 0.000000 , 160.754776]
    [0.000000 , 390.301514 , 135.610931]
    [0.000000 , 0.000000 , 1.000000]

M = [388.645844 , 0.000000 , 182.932190]
    [0.000000 , 389.123718 , 124.230911]
    [0.000000 , 0.000000 , 1.000000]

D = [-0.391242 , 0.173672 , -0.002223 , 0.001944]

D = [-0.396425 , 0.209652 , -0.000585 , -0.001177]
```

Ma trận F (fundamental matrix) của hệ camera sử dụng và bảng sai số đo ngẫu nhiên:

```
F = [-0.000008 , 0.000093 , -0.014982]
    [0.000107 , -0.000002 , 0.127069]
    [-0.007596 , -0.162804 , 1.000000]
```

Camera trái	Camera phải	Sai số
(21,182)	(144,158)	0.012
(79,64)	(142,51)	-0.18
(88,157)	(163,134)	0.2
(48,205)	(60,171)	0.009
(37,34)	(42,31)	-0.355
(31,122)	(43,103)	-0.24
...	...	...

Ma trận chuyển đổi hệ tọa độ giữa hai camera trai và phải

```
R = [0.869440 , 0.006914 , -0.493990]
    [-0.003767 , 0.999966 , 0.007366]
    [0.494024 , -0.004543 , 0.869436]
T = [265.918793 , 10.469414 , 65.164299]
```

Khoảng cách giữa hai camera: 273.986859 mm

Kết quả sai số khi tính khoảng cách đến một đối tượng cho trước (tham khảo [])

Khoảng cách truyền và tốc độ truyền (môi trường: trong nhà)

Khoảng cách (m)	Tốc độ (fps)	Nén dữ liệu trước khi truyền
1 - 3	22	Không
3 - 5	20	Không
	22	Có (JPEG)
5-8	15	Không
	22	Có (JPEG)
8 – 15	10	Không
	20	Có (JPEG)
> 15	5 – 7	Không
	15	Có (JPEG)

IV. KẾT LUẬN

Do thời gian thực hiện có hạn nên đề tài chỉ tập trung vào việc xây dựng mô hình hệ thống stereo (phần cứng), bộ phần mềm và tiến hành **khảo sát thực tế hệ thống** xây dựng dựa trên tiêu chuẩn là mục đích sử dụng (**trên mô hình robot di động**). Việc áp dụng các kết quả **trên robot di động trong thực tế** chưa hiện thực được. Tuy nhiên đây là những bước đầu tiên trong quá trình kết

hợp các kết quả nghiên cứu lý thuyết giữa lĩnh vực máy tính – thị giác (computer vision) và lĩnh vực robot. Kết quả của đề tài có thể được sử dụng hiệu quả hơn bằng cách sử dụng chúng vào trong các hướng nghiên cứu khác như nghiên cứu chuyển động của đối tượng (ví dụ bám mục tiêu di động, phân tích quỹ đạo chuyển động của vật cản,...).

TÀI LIỆU THAM KHẢO

[1] Vương Ngọc Dũng, *Khảo nghiệm phương pháp hiệu chuẩn camera được đề xuất bởi Zhengyou Zhang*, Hội nghị khoa học trẻ Bách khoa năm 2005.

[2] Vương Ngọc Dũng, Luận văn CH *Ứng dụng kỹ thuật xử lý ảnh trong điều khiển,* 2005

[3] Lương Mạnh Bá & Nguyễn Thanh Thủy, *Nhập môn xử lý ảnh số,* Nhà xuất bản Khoa Học và Kỹ Thuật, (1999)

[4] **R**afael C. Gonzalez & Richard E. Woods, *Digital Image Processing*, Addison – Wesley Publishing company, (1993)

[5] Zhengyou Zhang, *Flexible camera calibration by viewing a plane from unknown orientations*, Microsoft research, One microsoft way, Redmond, WA 98052-6399, USA 1998

[6] Course về computer vision (computer vision homepage: www.cs.cmu.edu/~cil/ vision.html).

[7] Luận văn đại học của một số sinh viên do tác giả hướng dẫn.