

PHÁT TRIỂN THUẬT TOÁN XỬ LÝ ẢNH ĐỂ PHÁT HIỆN VÀ ƯỚC LƯỢNG KHOẢNG CÁCH TỪ HỆ CAMERA ĐẾN TÂM QUẢ CÀ CHUA CHÍN TRÊN CÂY

Trương Quốc Bảo¹, Quách Tuấn Văn² và Nguyễn Minh Luân¹

¹ Khoa Công nghệ, Trường Đại học Cần Thơ

² Phòng Đào tạo, Trường Cao đẳng Y tế Cà Mau

Thông tin chung:

Ngày nhận: 01/11/2014

Ngày chấp nhận: 26/02/2015

Title:

Developing computer vision algorithm for ripe tomato localization and estimation of the distance from the camera system to the centre of the ripe tomato on the tree

Từ khóa:

Robot thu hoạch trái cây, phát hiện và nhận dạng quả cà chua chín, mặt phẳng kết tủa màu r-g, phân tích hình dáng, hiệu chỉnh camera

Keywords:

Fruit harvesting robot, ripe tomato detection and recognition, r-g chromaticity plane, shape analysis, camera calibration

ABSTRACT

An efficient method for locating fruit on the tree and estimating the distance from the center of the fruit to camera system is one of the major requirements for the fruit harvesting system. This paper proposes a computer vision algorithm for detecting ripe tomatoes and estimating the distance from the camera system to the ripe tomato on the tree using stereo vision system. The main steps of the algorithm include color segmentation, region labeling, size filtering, and boundary determination for ripe tomato candidate regions, shape feature extraction for localization of the ripe tomato on the tree and distance estimation using stereo vision system. The experiment was performed with 100 images in which there were a total of 244 ripe tomatoes. The accuracy of ripe tomato recognition based on r-g chromaticity color space and based on color analysis was 96.7% and 88.9% respectively.

TÓM TẮT

Một kỹ thuật hiệu quả để định vị trái trên cây và ước lượng khoảng cách từ hệ thống camera đến tâm của trái là yêu cầu chính cho robot thu hoạch trái cây. Nghiên cứu này đề xuất một giải thuật xử lý ảnh mới để nhận dạng và định vị quả cà chua chín trên cây đồng thời ước lượng khoảng cách từ hệ camera đến tâm của trái. Thuật toán bao gồm các bước chính: phân đoạn ảnh, gán nhãn, lọc kích thước, xác định đường biên cho các vùng ứng viên là quả cà chua chín, rút trích các đặc trưng hình dáng để định vị quả cà chua chín trên cây và cuối cùng là ước lượng khoảng cách sử dụng hệ 2 camera. Thử nghiệm được tiến hành trên tập dữ liệu 100 ảnh thực nghiệm với 244 quả cà chua chín cần nhận dạng. Độ chính xác của phương pháp được đề nghị là 96.7% đối với phương pháp nhận dạng dựa trên mặt phẳng kết tủa màu r-g và 88.9% đối với phương pháp phân tích màu sắc.

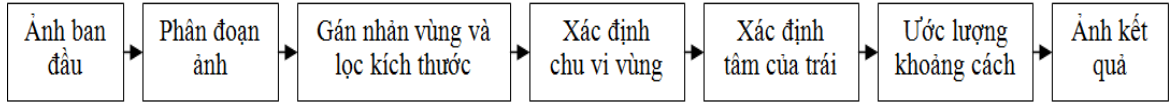
1 GIỚI THIỆU

Hệ thống nhận dạng và ước lượng vị trí tâm của trái trên cây sử dụng các kỹ thuật xử lý ảnh và ước lượng khoảng cách từ tâm của trái đến hệ thống quan sát là phần đầu tiên và quan trọng nhất trong

nghiên cứu, chế tạo robot hái lượm ứng dụng trong thu hoạch trái cây. Các hệ thống như vậy đang được phát triển và ứng dụng trong ngành nông nghiệp tự động hóa, thông minh ở một số quốc gia phát triển trên thế giới với nhiều công trình nghiên

cứu liên quan (Jian-jun Y. *et al*, 2009; Hannan M. W. *et al.*, 2009 và Hetal N.Patel, 2011). Tuy nhiên, các nghiên cứu như vậy ở Việt Nam là còn rất hạn chế.

Bài báo đề xuất một giải thuật mới để nhận dạng và định vị quả cà chua chín trên cây sử dụng kỹ thuật phân đoạn ảnh và phân tích hình dáng.



Hình 1: Quy trình xử lý của phương pháp được đề xuất

2 NỘI DUNG NGHIÊN CỨU

2.1 Hệ thống quan sát và thu nhận hình ảnh

Nhằm mục đích thu nhận hình ảnh quả cà chua chín, đưa về máy tính xử lý để nhận dạng, định vị, ước lượng vị trí tâm cũng như khoảng cách từ tâm quả cà chua chín được phát hiện trên cây đến hệ thống quan sát (Hình 2), mô phỏng thay cho cánh tay robot.



Hình 2: Hệ thống quan sát kết hợp giữa 2 cameras và máy tính

Trong hệ thống này, chúng tôi thiết kế một giá đỡ cơ khí để hỗ trợ việc nâng đỡ hệ camera và laptop. Phần thứ nhất là một giá đỡ có khả năng tăng giảm độ cao trong khoảng 50-150 cm và có thể xoay từ 0°-360° mô phỏng giống như một cách tay robot có thể chuyển đổi vị trí trong không gian với hai camera quan sát để xác định vị trí quả cà chua cần hái trên cây. Phần thứ hai của hệ thống là hệ camera gồm hai camera Logitech HD Pro C920 độ phân giải 15-Megapixel, mỗi camera có khả năng xử lý và xác định hình ảnh với khoảng cách gần nhất từ hệ camera đến vật cần đo khoảng cách là 10 cm. Cuối cùng là một laptop có cấu hình tương đối mạnh để xử lý ảnh một cách nhanh chóng.

2.2 Phân đoạn ảnh

Phân đoạn ảnh là công đoạn rất quan trọng trong các bài toán xử lý ảnh và thị giác máy tính. Trong nghiên cứu này, quá trình phân đoạn ảnh

Khoảng cách từ tâm quả cà chua chín đến hệ camera được ước lượng dựa trên kỹ thuật hiệu chỉnh camera và phương pháp đo khoảng cách do Manaf A. Mahammed *et al*, 2013 đề xuất. Quy trình xử lý của phương pháp được trình bày như Hình 1.

nhằm tách quả cà chua chín cần nhận dạng ra khỏi nền. Chúng tôi tiến hành so sánh 2 thuật toán phân đoạn ảnh để phát hiện quả cà chua chín là *phương pháp phân tích màu sắc và phân đoạn ảnh dựa trên mặt phẳng kết tủa màu r-g*.

2.2.1 Phương pháp phân tích màu sắc

a. Loại trừ nền

Từ nghiên cứu thực nghiệm, chúng tôi tiến hành xét cà chua chín là những quả có màu chuyển sang đỏ hơn 90%. Các quả có màu vàng – cam là những quả gần chín tùy theo yêu cầu thực tế có thể thu hoạch hoặc không. Quá trình loại trừ nền được thực hiện như sau:

Bước 1: Tách kênh màu R, G, B từ ảnh đầu vào để được các ảnh trên các kênh màu riêng biệt.

Bước 2: Lấy ảnh thành phần R trừ ảnh thành phần G.

Bước 3: Nếu giá trị R-G thỏa điều kiện lớn hơn một giá trị ngưỡng định trước thì các điểm ảnh tương ứng sẽ được giữ lại (Hình 3b).

Kết quả đạt được sẽ bao gồm những quả cà chua chín và gần chín. Để thu hoạch những quả thật sự chín (màu đỏ đậm), công đoạn kế tiếp cần thực hiện là loại bỏ những quả màu vàng – cam.

b. Xác định các quả cà chua thật sự chín

Từ ảnh đã được loại bỏ nền (Hình 3b), ta thực hiện:

Bước 1: Trích xuất các điểm ảnh màu vàng. Ảnh sau khi trừ nền lần lượt được chuyển đổi sang không gian màu HSI và YIQ. Các điểm ảnh màu vàng sẽ được trích xuất theo công thức (1):

$$P_{yellow} = 1.2 * Y - S \text{ với } Y \text{ là độ sáng, } S \text{ là độ bão hòa. (1)}$$

Bước 2: Tiếp tục trích xuất các điểm ảnh màu đỏ và cam theo công thức (2):

$$P_{red_orange} = R - 1.3 * G \quad (2)$$

Bước 3: Trích xuất một số điểm ảnh màu đỏ đại diện để xác định quả cà chua thật sự chín. Các điểm ảnh màu đỏ được trích xuất theo điều kiện (3):

$$\begin{cases} P_{red} = I - Y < 0.01 \\ Q < 0.08 \end{cases} \quad \text{với } Y \text{ là độ sáng, } I \text{ là}$$

màu sắc, Q là độ bão hòa. (3)

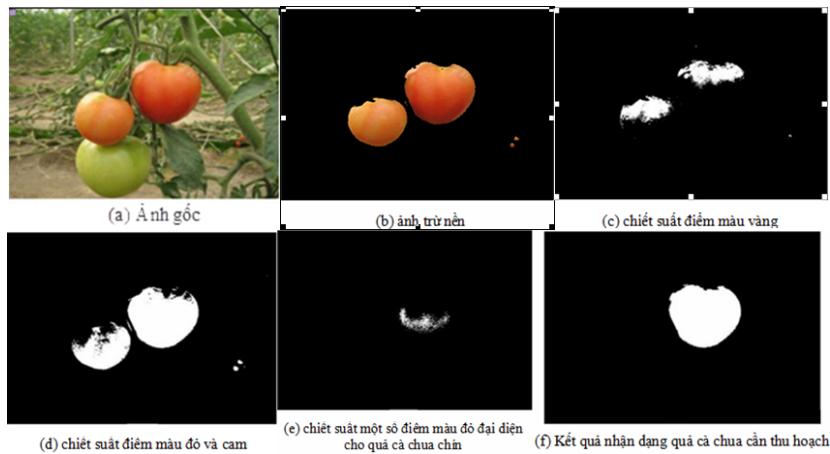
Bước 4: Sau khi thực hiện các bước từ 1 đến 3 ta được các ảnh kết quả bao gồm các điểm ảnh vàng, đỏ, cam. Các ảnh này sẽ được chuyển sang ảnh nhị phân dựa trên giá trị ngưỡng thực nghiệm.

Các điểm ảnh còn lại của một vùng sẽ đại diện cho quả cà chua chín tương ứng. Các vùng không có điểm đại diện đồng nghĩa với việc các quả cà chua tương ứng sẽ bị loại bỏ.

Bước 5: Sử dụng phương pháp 8 lân cận tái cấu trúc (Luc Vincent, 1993) ảnh 3e từ ảnh kết quả 3d để loại bỏ quả cà chua màu cam trong ảnh 3d.

Bước 6: Thêm lại phần màu vàng cho quả cà chua chín bằng cách lấy ảnh kết quả của bước 5 cộng với ảnh 3c.

Bước 7: Từ ảnh kết quả bước 6, tiếp tục tái cấu trúc thêm lần nữa, ta sẽ nhận được quả cà chua cần xác định để thu hoạch (Hình 3f).



Hình 3: Kết quả nhận dạng quả cà chua cần thu hoạch

2.2.2 Phương pháp phân đoạn ảnh dựa trên mặt phẳng kết tủa màu r-g

Trong thực tế, quả cà chua chín cần được ra khỏi nền bằng một kỹ thuật phân đoạn ảnh. Nếu chúng ta sử dụng không gian màu RGB thì cần phân tích trên từng kênh màu riêng biệt sẽ làm giảm hiệu suất và tăng thời gian xử lý chung của cả hệ thống như phương pháp phân tích màu sắc trên. Việc sử dụng kỹ thuật phân đoạn ảnh dựa trên mặt phẳng kết tủa màu r-g là phù hợp hơn.

a. Chuyển đổi từ không gian màu RGB sang mặt phẳng kết tủa màu r-g

Hình 4 mô tả mô hình chuyển đổi tọa độ điểm ảnh từ không gian màu RGB sang mặt phẳng kết tủa màu r-g (Balkenius C. and Johansson B., 2007). Một điểm ảnh trong không gian RGB được chuyển đổi thành màu (r, g) trong mặt phẳng kết tủa màu r-g theo công thức (4):

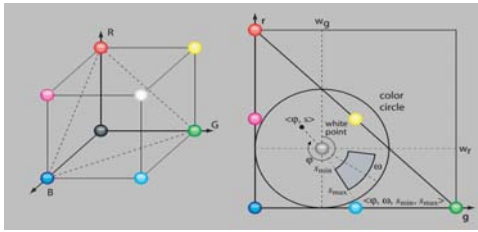
$$r = \frac{R}{R+G+B}; \quad g = \frac{G}{R+G+B} \quad (4)$$

Mặt phẳng kết tủa màu r-g chứa các màu của không gian RGB ban đầu nhưng cường độ màu của các điểm ảnh đã được loại bỏ. Cường độ màu có thể được tính lại theo (5):

$$I = R + G + B \quad (5)$$

Phép biến đổi trên ánh xạ các giá trị màu R, G, B lên 3 đỉnh của tam giác màu. Màu xám và trắng được ánh xạ tại cùng một điểm $(\frac{1}{3}, \frac{1}{3})$, gọi là điểm trắng của tam giác màu.

Điểm trắng của ảnh có thể được ước lượng từ dữ liệu hình ảnh nếu chúng ta giả sử rằng trung bình màu của ảnh là xám. Giá trị trung bình của các kênh màu, được tính theo (6):



Hình 4: Chuyển đổi không gian màu từ RGB sang r-g (Balkenius C. and Johansson B., 2007)

$$\bar{r} = \frac{\bar{R}}{\bar{R} + \bar{G} + \bar{B}}; \quad \bar{g} = \frac{\bar{G}}{\bar{R} + \bar{G} + \bar{B}} \quad (6)$$

Điểm trắng được đặt tại vị trí: $w = \langle \bar{r}, \bar{g} \rangle$. Mỗi màu được biểu diễn bằng 2 giá trị: (1) góc xung quanh điểm trắng j và (2) giá trị bão hòa màu s là khoảng cách từ điểm trắng được tính theo (7):

$$j = \arctan\left(\frac{r - \bar{r}}{g - \bar{g}}\right); \quad s = \sqrt{(r - \bar{r})^2 + (g - \bar{g})^2} \quad (7)$$

b. Phân đoạn ảnh sử dụng không gian kết tủa màu r-g

Để tìm một vùng có màu xác định nào đó của ảnh, đầu tiên ta cần phải phân loại các điểm ảnh. Một khu vực màu được xác định bằng một bộ bốn giá trị: $\langle j, s_{\min}, s_{\max}, w \rangle$ với j là góc xung quanh điểm trắng, s_{\min}, s_{\max} lần lượt là giá trị nhỏ nhất và lớn nhất của độ bão hòa màu, w là độ rộng của vùng màu đang xét. Quá trình phân đoạn ảnh được thực hiện bằng cách tìm tất cả các điểm ảnh trong mặt phẳng r-g có giá trị s và j thỏa các ngưỡng, s_{\min}, s_{\max} và j_t được xác định theo (8):

$$\begin{cases} s_{\min} < s < s_{\max} \\ j_t - \frac{w}{2} < j < j_t + \frac{w}{2} \end{cases} \quad (8)$$

Phương pháp phân đoạn ảnh dựa trên mặt phẳng kết tủa màu r-g cho phép ta thiết lập ngưỡng màu để nhận dạng đối tượng một cách linh hoạt bằng cách thiết lập các giá trị ngưỡng cho bộ bốn $\langle j, s_{\min}, s_{\max}, w \rangle$. Trong thực tế, đôi khi ta muốn hái những quả thật chín (đỏ đậm) hoặc một số quả vừa chín tới (đỏ nhạt hay đỏ ngả vàng) (Hình 5).

2.3 Thuật toán ước lượng vị trí tâm quả cà chua chín trên cây

Sau khi lọc bỏ nền, bước kế tiếp ta sẽ tiến hành gán nhãn, lọc kích thước, xác định đường biên cho các vùng ứng viên là quả cà chua chín được thực hiện theo quy trình sau.



Hình 5: Kết quả phân đoạn ảnh với các bộ giá trị khác nhau của $\langle j, s_{\min}, s_{\max}, w \rangle$. Từ trái qua:

ảnh đầu vào; ảnh phân đoạn với bộ giá trị $\langle j = 85^\circ, s_{\min} = 0.154398, s_{\max} = 0.745356, w = 65^\circ \rangle$;

ảnh phân đoạn với bộ giá trị $\langle j = 105^\circ, s_{\min} = 0.154398, s_{\max} = 0.745356, w = 50^\circ \rangle$

2.4 Thuật toán ước lượng vị trí tâm quả cà chua chín trên cây

Sau khi lọc bỏ nền, bước kế tiếp ta sẽ tiến hành gán nhãn, lọc kích thước, xác định đường biên cho các vùng ứng viên là quả cà chua chín được thực hiện theo quy trình sau.

2.4.1 Gán nhãn, lọc kích thước, xác định đường biên của các vùng ứng viên

a. Gán nhãn

Mục đích của thao tác này là đánh dấu các vùng liên thông trên ảnh. Chúng tôi sử dụng phương pháp chập mặt nạ đơn giản (Alain Boucher, 2012) để đánh dấu các vùng liên thông trên ảnh và gán nhãn cho tất cả các vùng liên thông là vùng ứng viên cho quả cà chua.

b. Lọc kích thước

Trong các vùng ứng viên, có những vùng kích thước quá lớn hoặc quá nhỏ sẽ không thể là quả cà chua. Ngoài ra, từ thực tế cho thấy tùy thuộc vào khoảng cách từ camera đến cây cà chua mà kích thước của các quả cà trên ảnh sẽ thay đổi. Vì vậy, với mỗi vùng được gán nhãn, ta tính tổng số điểm ảnh của vùng đó, nếu nhỏ hơn độ lớn của bộ lọc kích thước `size_filter` thì loại bỏ vùng này.

c. Xác định đường biên

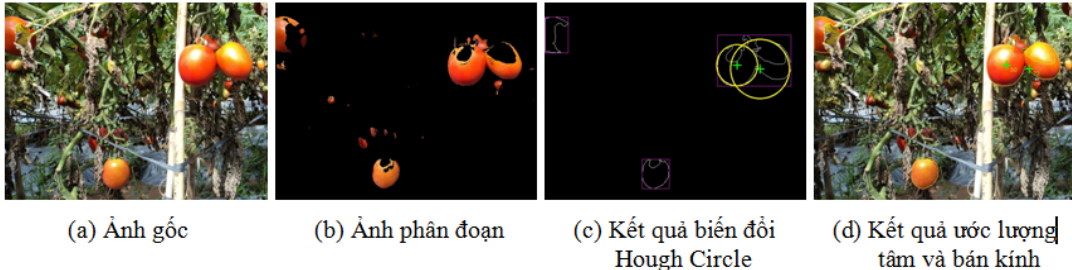
Quá trình trích lọc chu vi tương tự như quá trình dò tìm đường biên theo phương pháp 4 lân cận. Một điểm nằm hoàn toàn trong vùng chu vi sẽ có 4 điểm lân cận thuộc vùng nhãn, một điểm nằm

trên đường chu vi sẽ có ít nhất 1 điểm lân cận thuộc vùng ảnh nền (Suzuki S. and Abe K., 1985).

2.4.2 Định vị vị trí quả cà chua chín sử dụng phương pháp biến đổi Hough Circle

Phép biến đổi Hough Circle cho phép phát hiện các vật thể có dạng tương tự hình tròn và xác định vị trí tâm, độ dài bán kính của đường tròn mô tả

cho các vật thể này (Yuen H. K. et al., 1990). Cách tiếp cận này đơn giản, nhưng kết quả thu được đôi khi không chính xác. Trong trường hợp một nhóm nhiều quả cà chua chồng lấp lên nhau, Hough Circle có thể nhận dạng thừa hoặc thiếu một vài quả hoặc xác định vị trí tâm không phù hợp (Hình 6).



Hình 6: Hough Circle phát hiện không đúng số quả và ước lượng tâm không phù hợp

Để khắc phục nhược điểm của phương pháp phân tích Hough Circle, chúng tôi đề nghị một phương pháp ước lượng vị trí tâm và độ dài bán kính dựa trên ý tưởng của Hannan M. W. et al, 2009 gọi là kỹ thuật phân tích hình dáng của đường biên (contour) để xây dựng đường tròn đại diện cho vị trí quả cà chua chín.

2.4.3 Định vị vị trí quả cà chua chín sử dụng phương pháp phân tích hình dáng

Phương pháp phân tích hình dáng cho việc phát hiện và xây dựng đường tròn đại diện cho vị trí của quả cà chua bao gồm 2 giai đoạn là tính toán các tham số đường tròn cho các cung ứng viên trên đường chu vi và ước lượng vị trí tâm cùng độ dài bán kính của đường tròn đại diện.

a. Tính toán các tham số đường tròn

Bắt đầu từ một vị trí bất kỳ ta trượt dọc trên đường chu vi theo từng phân đoạn (mỗi phân đoạn

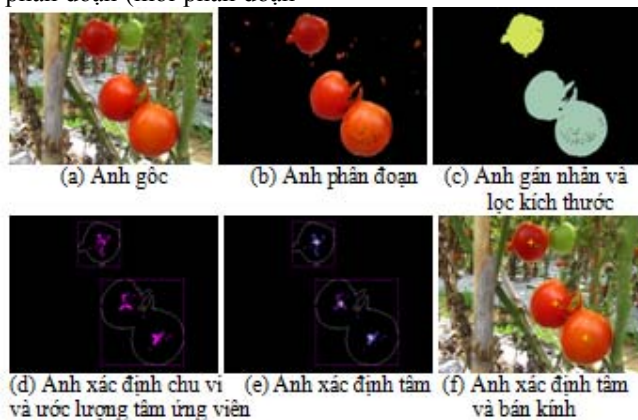
50 điểm ảnh). Với mỗi phân đoạn, chọn 3 điểm ảnh: đầu, cuối và giữa làm dữ liệu vào thành lập hệ:

$$(x_i - x_c)^2 + (y_i - y_c)^2 = r_c^2 \quad (i=1..3) \quad (9)$$

(x_i, y_i) là điểm ảnh thứ i thuộc phân đoạn đang xét; (x_c, y_c) là tọa độ tâm; r_c là bán kính đường tròn.

b. Ước lượng vị trí tâm, độ dài bán kính của đường tròn đại diện cho vị trí quả cà chua

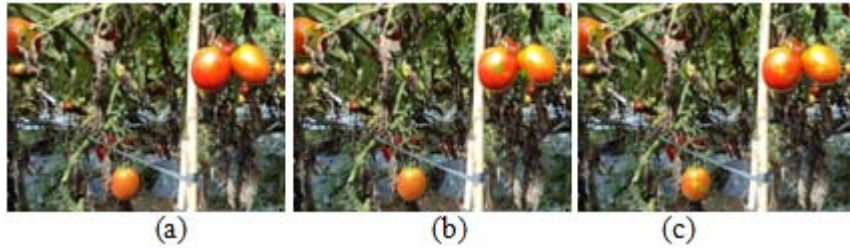
Tiếp theo ta tiến hành gom các tâm ứng viên này thành từng nhóm dựa vào khoảng cách của chúng. Các tâm lân cận nhau sẽ được gom thành một nhóm. Nếu nhóm tâm chứa ít hơn 3 tâm sẽ được loại bỏ. Ngược lại, ta tìm vị trí trung tâm của nhóm bằng cách tính trung bình giá trị các tọa độ của các tâm ứng viên trong cùng một nhóm. Vị trí trung tâm này được xem là tâm của đường tròn đại diện cho vị trí quả cà chua (Hình 7).



Hình 7: Quá trình nhận dạng và định vị vị trí quả cà chua chín trên cây

Tiếp cận phân tích hình dáng do chúng tôi đề xuất có thể khắc phục các điểm yếu của biến đổi

Hough Circle như phát hiện thừa, thiếu tâm hay ước lượng vị trí tâm không phù hợp (Hình 8)



Hình 8: Ảnh gốc (a); Phương pháp Hough Circle (b); Phương pháp phân tích hình dáng (c)

So với phương pháp được đề nghị bởi Hannan M. W. *et al.* 2009, cách tiếp cận của chúng tôi có 2 điểm khác biệt: (1) Sử dụng phân đoạn 50 điểm ảnh và (2) mỗi phân đoạn sử dụng 3 điểm ảnh đầu, giữa và cuối của mỗi phân đoạn để xây dựng hệ phương trình tìm các tham số đường tròn. Vì mỗi đường tròn cần 3 tham số: tọa độ tâm (x_c, y_c) và bán kính r_c nên chỉ cần hệ 3 phương trình là đủ xác định các tham số một cách chính xác. Hannan M. W. *et al.* (2009) sử dụng phương pháp bình phương nhỏ nhất để xác định gần đúng giá trị các thông số của đường tròn nên để tăng độ chính xác phải sử dụng phân đoạn với nhiều điểm ảnh (100 điểm ảnh) điều này làm tăng thời gian xử lý lên rất nhiều. Phương pháp cải tiến của chúng tôi tìm giá trị chính xác của các thông số thay cho giá trị gần đúng. Ngoài ra, độ chính xác đối với phương pháp do chúng tôi đề nghị sẽ càng tăng khi độ dài phân đoạn càng nhỏ. Do đó, để cân đối chúng tôi chọn mỗi phân đoạn 50 điểm ảnh. Qua quá trình thực nghiệm phương pháp cải tiến mất trung bình 0.18 giây/ảnh để xác định tâm và bán kính so với phương pháp của Hannan M. W. *et al.* 2009 là khoảng 0.34 giây/ảnh.

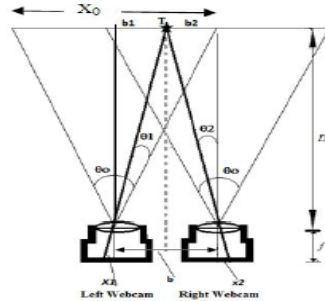
2.5 Ước lượng khoảng cách từ hệ camera quan sát đến vị trí tâm quả cà chua chín trên cây

Một số phương pháp thường sử dụng để ước lượng khoảng cách từ hệ camera quan sát đến vật thể như sử dụng sóng siêu âm, sóng laser, sóng hồng ngoại... Trong nghiên cứu này, chúng tôi sử dụng mô hình 2 cameras như mô tả ở Hình 9 được sử dụng phổ biến cho các hệ điều khiển robot tự động, kết hợp cùng các thuật toán xử lý ảnh như hiệu chỉnh hệ camera và áp dụng kỹ thuật được trình bày bởi Manaf A. Mahammed *et al.* 2013 để ước lượng khoảng cách từ hệ camera đến tâm quả cà chua chín trên cây.

2.5.1 Hiệu chỉnh camera (Camera calibration)

Quá trình chuyển đổi ảnh từ không gian thực lên không gian hình ảnh thường cho kết quả không tối ưu vì ảnh thu được từ hệ 2 cameras là không

đồng nhất và thường bị méo (distortion). Nếu chúng ta không có thông tin về mối liên hệ giữa 2 cameras, khi đó nếu chúng ta muốn tìm một điểm tương ứng trên ảnh bên trái ở ảnh bên phải chúng ta phải quét toàn bộ ảnh bên phải.



Hình 9: Sơ đồ nguyên lý được sử dụng để ước lượng khoảng cách từ hệ camera đến vật thể (Manaf A. Mahammed *et al.*, 2013)

Tuy nhiên, nếu ta có được mối liên hệ giữa 2 cameras thì việc tìm một điểm ảnh thuộc ảnh bên phải tương ứng với ảnh bên trái có thể được thực hiện bằng cách dò tìm trên 1 đường cong epipolar duy nhất (R. Hartley and A. Zisserman, 2000).

Ngoài ra, nếu cả 2 cameras được đặt song song, đồng trục với nhau trên mặt phẳng và có cùng các thông số về tiêu cự, độ méo của ảnh,... khi đó các đường cong epipolar sẽ được hiệu chỉnh trở thành các đường thẳng. Điều đó đưa đến một hệ quả quan trọng là điểm ảnh thuộc ảnh bên trái sẽ nằm trên cùng một đường thẳng trên ảnh bên phải. Kỹ thuật được sử dụng để thực hiện việc này gọi là hiệu chỉnh cameras. Thao tác hiệu chỉnh cameras thường được thực hiện bằng cách chụp các mẫu bản cờ ở nhiều góc độ khác nhau để tính ra các thông số nội của camera. Dựa vào các thông số này chúng ta có thể hiệu chỉnh lại độ méo của ảnh bằng một giải thuật phù hợp (R. Hartley and A. Zisserman, 2000), đồng thời hiệu chỉnh độ méo của ảnh để biến các đường cong epipolar thành các đường thẳng (Hình 10).



(a) ảnh trước khi hiệu chỉnh cameras

(b) ảnh sau khi hiệu chỉnh cameras

Hình 10: Ảnh kết quả so sánh trước (a) và sau khi hiệu chỉnh cameras (b)

2.5.2 Ước lượng khoảng cách từ hệ camera đến tâm quả cà chua chín trên cây

a. Phương trình disparity cơ bản

Mô hình 2 cameras quan sát và đo khoảng cách được trình bày như ở Hình 9. Trong đó, hệ 2 cameras được đặt đồng trục trên cùng một mặt phẳng cách nhau một khoảng cách là b . Hệ cameras được hiệu chỉnh để có cùng bộ thông số nội, cùng tiêu cự f và góc nhìn θ_0 . Ảnh đối tượng quan sát được T xuất hiện ở hoành độ x_1 trên ảnh bên trái và ở hoành độ x_2 trên ảnh bên phải. Áp dụng nguyên lý đồng dạng ta có:

$$\frac{b_1}{D} = -\frac{x_1}{f}, \quad \frac{b_2}{D} = \frac{x_2}{f} \tag{10}$$

Vì $b = b_1 + b_2$ nên: $b = \frac{D}{f}(x_2 - x_1)$ hay

$$D = \frac{Df}{(x_2 - x_1)} \tag{11}$$

Ngoài ra: $\tan\left(\frac{\theta_0}{2}\right) = \frac{x_0}{D} = \frac{x_1}{f}$ (12)

Từ các phương trình (12) và (13) ta có khoảng cách D từ hệ camera đến vật thể:

$$D = \frac{bx_0}{2 \tan\left(\frac{\theta_0}{2}\right)(x_2 - x_1)} \tag{13}$$

Với x_0 là chiều rộng ảnh, $x_2 - x_1$ là sự khác biệt (disparity) về vị trí của đối tượng trên ảnh trái và phải cùng tính theo pixels. Theo Manaf A. Mohammed *et al*, 2013 khoảng cách D sẽ tỷ lệ nghịch với hiệu $(x_2 - x_1)$ vì vậy để bù cho lỗi cho góc quan sát θ_0 thì θ_0 sẽ được cộng thêm một đại lượng ϕ trong thành phần tan của biểu thức (13) và được xác định bằng thực nghiệm.

$$D = \frac{bx_0}{2 \tan\left(\frac{\theta_0}{2} + \phi\right)(x_2 - x_1)} \tag{14}$$

Từ đây biểu thức (14) có thể được biểu diễn về dạng hàm mũ như biểu thức (15):

$$D = k * x^z \tag{15}$$

Trong đó z là hằng số được xác định bằng thực nghiệm, k và x được xác định bởi (16):

$$k = \frac{bx_0}{2 \tan\left(\frac{\theta_0}{2} + \phi\right)} \quad \text{và} \quad x = x_2 - x_1 \tag{16}$$

b. Xác định khoảng cách từ hệ cameras đến vật thể

Để tìm các giá trị k và z trong phương trình (15) giá trị độ chênh lệch x sẽ được xác định với khoảng giá trị khoảng cách D từ hệ camera đến vật thể được xác định trước. Chúng tôi tiến hành thực nghiệm đo các bộ giá trị (x, D) với D từ 15 cm đến 200 cm và phân tích hồi quy, tìm được:

$$D(\text{cm}) = 5926 * x(\text{pixel})^{-1.005} \tag{17}$$

Áp dụng (17) để xác định khoảng cách từ hệ camera đến vật thể chúng tôi nhận được kết quả rất tốt với sai số đo được là nhỏ hơn 1 cm và hệ số tương quan hồi quy là $R^2 = 0.9996$.

c. Áp dụng ước lượng khoảng cách từ hệ cameras đến tâm quả cà chua chín trên cây

Để ước lượng khoảng cách từ hệ camera quan sát đến tâm các quả cà chua chín được xác định, chúng tôi tiến hành áp dụng thuật toán định vị vị trí tâm của các quả cà chua chín cho ảnh bên trái, sau đó tiến hành tìm kiếm các tâm tương ứng của các quả cà chua chín trên ảnh bên phải. Phương pháp sử dụng được gọi là tìm tổng bình phương sự khác biệt về mức xám giữa các pixels (Sum of Square Differences – SSD). Một mặt nạ 3x3 được quét qua tất cả các điểm ảnh ứng viên trên ảnh bên phải và điểm ảnh trên ảnh bên phải cho giá trị độ khác biệt nhỏ nhất là điểm ảnh cần tìm. Thuật toán được mô tả như sau:

for mỗi pixel là tâm của quả cà trên ảnh bên trái do {

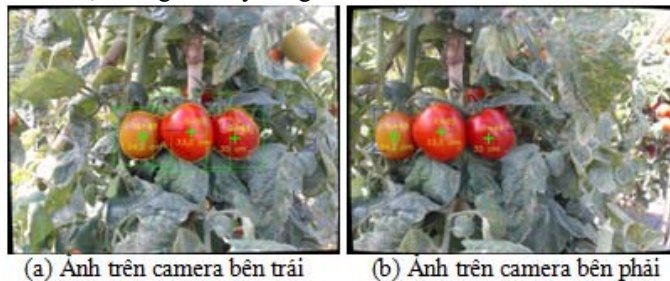
Trượt mặt nạ lên các pixels ứng viên thuộc đường thẳng tương ứng trên ảnh bên phải.

Tính giá trị Cost cho từng pixel ứng viên theo công thức (18):

$$Cost(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 [f_L(x+i, y+j) - f_R(x+i, y+j)]^2 \quad (18)$$

return (pixel(x, y)) tương ứng với giá trị của Cost(x, y) là nhỏ nhất.

Hình 11 trình bày một kết quả thực nghiệm của hệ thống. Từ kết quả Hình 11, chúng ta thấy rằng



Hình 11: Kết quả ước lượng khoảng cách từ hệ camera đến tâm quả cà chua chín trên cây

3 KẾT QUẢ VÀ THẢO LUẬN

Để đánh giá hiệu quả của phương pháp đề nghị, chúng tôi tiến hành chụp và sưu tập dữ liệu ảnh cà chua từ các vườn cà chua thuộc phạm vi thành phố Đà Lạt. Các ảnh thu được sẽ được tiến hành tiền xử lý về dạng ảnh JPEG với độ phân giải 640x480. Tập dữ liệu bao gồm 100 ảnh với 244 quả cà chua chín cần nhận dạng. Để dễ dàng phân tích chúng tôi chia tập dữ liệu ảnh thành 8 tập con:

- Tập 1 gồm những ảnh không chứa quả cà chua chín cần nhận dạng nào,
- Tập 2 gồm những ảnh chỉ chứa đúng 1 quả cà chua chín cần nhận dạng,

quá trình xác định vị trí tâm quả cà trên ảnh camera phải hoàn toàn phù hợp với vị trí tâm quả cà được ước lượng trên ảnh camera trái. Qua quá trình **26 lần thực nghiệm đo khoảng cách trên các ảnh có nhiều quả cà chua chín cần nhận dạng. Tổng số quả cà chua cần nhận dạng là 56 quả**, kết quả đo khoảng cách từ hệ camera quan sát đến vị trí tâm quả cà cần xác định trên cây là rất khả quan với sai số trung bình khi ước lượng là **0.41 cm**. Việc ước lượng vị trí tâm và đo khoảng cách được thực hiện hoàn toàn online và không bị gián đoạn dựa trên hệ thống cơ khí quan sát đã được xây dựng.

- Tập ảnh thứ i gồm những ảnh chứa đúng i - 1 quả cà chua chín cần nhận dạng (i=3..8).

Chúng tôi tiến hành thực nghiệm nhận dạng quả cà chua chín trên các tập dữ liệu ảnh đã phân chia bằng 3 phương pháp: (1) kỹ thuật phân tích hình dáng dựa trên không gian kết tủa màu r-g đã đề nghị; (2) kỹ thuật nhận dạng dựa trên sự kết hợp phân tích màu sắc và hình dáng; (3) biến đổi Hough Circle dựa trên không gian kết tủa màu r-g. Đối với mỗi phương pháp chúng tôi thống kê số lượng quả cà chua chín được phát hiện trên từng ảnh, số lượng quả nhận dạng đúng. Kết quả được trình bày như trong Bảng 1.

Bảng 1: Thống kê kết quả nhận dạng của ba phương pháp đối với từng tập dữ liệu ảnh

Tập ảnh		1	2	3	4	5	6	7	8
Số lượng ảnh		11	16	28	24	10	7	3	1
Số lượng quả cần nhận dạng		0	16	56	72	40	35	18	7
Số quả đã nhận dạng	Phương pháp (1)	0	16	57	72	42	33	18	7
	Phương pháp (2)	0	16	53	65	39	29	14	5
	Phương pháp biến đổi Hough Circle (3)	0	16	53	54	27	23	10	4
Số quả nhận dạng đúng	Phương pháp (1)	0	16	56	70	38	32	17	7
	Phương pháp (2)	0	15	54	64	37	28	14	5
	Phương pháp biến đổi Hough Circle (3)	0	16	51	48	25	21	10	4

Từ thống kê ở Bảng 1 ta thấy trong tất cả các trường hợp, đặc biệt, đối với ảnh càng phức tạp (số quả cần nhận dạng nhiều) thì sự vượt trội của 2 phương pháp được đề nghị càng thể hiện rõ so với

số quả nhận dạng đúng nhiều hơn hẳn so với phương pháp biến đổi Hough Circle. Ngoài ra phương pháp nhận dạng dựa trên không gian kết

tủa màu r-g cũng cho kết quả hứa hẹn hơn so với phương pháp phân tích màu sắc.

Kết quả so sánh tổng hợp hiệu quả của 3 phương pháp được trình bày như trong Bảng 2.

Độ chính xác tổng hợp của quá trình nhận dạng theo các phương pháp phân tích hình dáng dựa trên các không gian màu khác nhau được đề nghị có độ chính xác lần lượt là 96.7% (phương pháp 1) và 88.9% (phương pháp 2) vượt trội hơn hẳn so với phương pháp biến đổi Hough Circle chỉ là 71.7%. Tóm lại, các kết quả phân tích và so sánh ở Bảng 1 và Bảng 2 chỉ ra rằng phương pháp phân tích hình dáng do chúng tôi đề nghị tốt hơn hẳn hướng tiếp cận theo phép biến đổi Hough Circle và hướng tiếp cận dựa trên mặt phẳng kết tủa màu r-g cho kết quả tốt nhất.

Bảng 2: Kết quả so sánh tổng hợp độ chính xác của 3 phương pháp nhận dạng

Số lượng ảnh		100
Số lượng quả cần nhận dạng		244
Phương pháp (1)	Số quả đã nhận dạng	245
	Số quả nhận dạng đúng	236
	Độ chính xác (%)	96.7
Phương pháp (2)	Số quả đã nhận dạng	221
	Số quả nhận dạng đúng	217
	Độ chính xác (%)	88.9
Phương pháp Hough Circle (3)	Số quả đã nhận dạng	187
	Số quả nhận dạng đúng	175
	Độ chính xác (%)	71.7

4 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

4.1 Kết luận

Trong nghiên cứu này, chúng tôi đã đề xuất một giải thuật xử lý ảnh hiệu quả để phát hiện, nhận dạng và ước lượng vị trí của quả cà chua chín trên cây cũng như khoảng cách từ tâm quả cà chua chín được xác định đến hệ thống quan sát. Kết quả của nghiên cứu này có thể là tiền đề cho việc thiết kế robot hái trái cà chua chín tự động. Các kết quả thực nghiệm và so sánh độ chính xác của phương pháp cho thấy phương pháp mặt phẳng kết tủa màu r-g kết hợp với kỹ thuật phân tích hình dáng là cho kết quả cao nhất (96.7%). Vì vậy, có thể áp dụng cho bài toán nhận dạng, định vị vị trí và ước lượng khoảng cách từ hệ camera quan sát đến quả cà chua chín trên cây.

4.2 Hướng phát triển

Trong tương lai, chúng tôi sẽ tiến hành nghiên cứu, phân tích, so sánh phương pháp đề xuất, cải tiến giải thuật để giảm thời gian xử lý cũng như

tăng hiệu quả nhận dạng. Đồng thời tiến hành thiết kế và thử nghiệm mô hình robot tự động hái quả cà chua chín trên cây và có mở rộng phương pháp để áp dụng cho các loại trái cây khác.

TÀI LIỆU THAM KHẢO

- Jian-jun Y., Han-ping M., Su-yu H., 2009. Segmentation Methods of Fruit Image based on Color Difference. Journal of Communication and Computer 6 (7): 40-45.
- Hannan M. W., Burks T. F., Bulanon D. M., 2009. A Machine Vision Algorithm Combining Adaptive Segmentation and Shape Analysis for Orange Fruit Detection. Agricultural Engineering International 11 (1): 1-17.
- Hetal N.Patel., Jain R. K., Joshi M. V., 2011. Fruit Detection using Improved Multiple Features based Algorithm. International Journal of Computer Applications 13 (2): 1-5.
- Manaf A. Mahammed, Amara I. Melhum, Faris A. Kochery., 2013. Object Distance Measurement by Stereo Vision. International Journal of Science and Applied Information Technology (IJSAIT) 2 (2): 5-8.
- Arman Arefi, Asad Modarres Motlagh, Rahman Farrokhi Teimourlou, 2010. A segmentation algorithm for the automatic recognition of tomato at harvest. Journal of Food, Agriculture & Environment 8 (3&4): 815-819.
- Suzuki S., Abe K., 1985. Topological Structural Analysis of Digitized Binary Image by Border Following. Computer Vision, Graphic, and Image Processing 30 (1): 32-46.
- Yuen H. K., Princen J., Illingworth J., Kittler J., 1990. Comparative Study of Hough Transform Methods for Circle Finding. Image Vision Computer 8 (1): 71-77.
- Luc Vincent, 1993. Morphological Grayscale Reconstruction. In Image Analysis: Applications and Efficient Algorithms. IEEE Transactions on Image Processing 2 (2): 176-201.
- Balkenius C., Johansson B., 2007. Finding Colored Objects in a Scene. CiteSeer.
- R. Hartley and A. Zisserman, 2000. Multiple View Geometry in Computer Vision. Cambridge University Press, 138-183.
- Alain Boucher, IFI 2012. Image processing and Computervision, slide 9 of course, Can Tho University.