



NGHIÊN CỨU XÁC ĐỊNH VỊ TRÍ VÀ GÓC NGHIÊNG CAMERA CHO ĐIỀU KHIỂN CÁNH TAY ROBOT CÔNG NGHIỆP SỬ DỤNG THỊ GIÁC ĐƠN TRÊN NỀN TẢNG ROS

Đoàn Công Luận

¹Trường Đại học Mỏ - Địa chất, 18 Phố Viên, Đông Ngạc, Hà Nội, Việt Nam

THÔNG TIN BÀI BÁO

CHUYÊN MỤC: Công trình khoa học

Ngày nhận bài: 28/10/2025

Ngày nhận bài sửa: 21/11/2025

Ngày chấp nhận đăng: 28/11/2025

Tác giả liên hệ:

Email: doancongluan@humg.edu.vn

TÓM TẮT

Bài báo này đề xuất giải pháp thị giác máy đơn (Monocular Vision) tích hợp trên robot công nghiệp nhằm xác định vị trí và tư thế không gian của vật thể trong môi trường làm việc phi cấu trúc. Thách thức lớn nhất đối với camera 2D là sai số do biến dạng phối cảnh khi trục quang học không vuông góc với bề mặt vật thể. Để giải quyết vấn đề này, nghiên cứu xây dựng thuật toán dựa trên ma trận đồng nhất (Homography) để ước lượng chính xác góc nghiêng (Tilt Angle) giữa trục camera và pháp tuyến mặt phẳng, đồng thời tính toán khoảng cách và độ lệch ngang (Lateral Offset) dựa trên mô hình Pinhole. Hệ thống được triển khai trên nền tảng ROS (Robot Operating System) với kiến trúc phân tán Publisher/Subscriber, cho phép tách biệt luồng xử lý ảnh và luồng điều khiển động học. Kết quả thực nghiệm trên robot Niryo Ned 2 với các vật thể hình học cơ bản cho thấy hệ thống hoạt động ổn định theo thời gian thực (30Hz), sai số ước lượng khoảng cách dưới 2% và xác định chính xác góc nghiêng để robot thực hiện bù trừ tư thế tiếp cận. Giải pháp này mang lại hiệu quả cao về chi phí và khả năng tích hợp cho các ứng dụng tự động hóa trong công nghiệp mỏ.

Từ khóa: ROS, thị giác máy, ước lượng góc lệch, Robot công nghiệp, Visual servoing.

@ Hội Khoa học và Công nghệ Mỏ Việt Nam

1. ĐẶT VẤN ĐỀ

Sự phát triển của thế hệ robot thông minh (Intelligent Robots) đang đánh dấu bước chuyển dịch quan trọng từ các dây chuyền sản xuất cố định sang các môi trường làm việc phi cấu trúc và linh hoạt. Trong bối cảnh này, khả năng tự động nhận dạng và thao tác với vật thể (object manipulation) dựa trên các cảm biến là yêu cầu cốt lõi. Khác với các robot công nghiệp truyền thống chỉ thực hiện chu trình lặp lại tại các tọa độ được lập trình sẵn, việc tích hợp thị giác máy vào vòng điều khiển của robot thông minh giúp robot thích ứng linh hoạt với sự thay đổi vị trí và tư thế ngẫu nhiên của các đối tượng thao tác [1].

Để giải quyết bài toán tri giác cho robot, các hệ thống camera 3D hoặc cảm biến độ sâu (Depth sensor) thường được xem là giải pháp tối ưu về độ

chính xác không gian. Tuy nhiên theo các đánh giá tổng quan, rào cản lớn nhất của các thiết bị này là chi phí đầu tư cao và yêu cầu tài nguyên tính toán lớn để xử lý lượng dữ liệu lớn, gây khó khăn cho việc triển khai đại trà trên các robot dịch vụ hoặc robot công nghiệp cỡ nhỏ [2, 3]. Ngược lại, giải pháp sử dụng Camera đơn 2D (Monocular Camera) mang lại lợi thế vượt trội về chi phí thấp, tính nhỏ gọn và khả năng tích hợp dễ dàng.

Tuy nhiên, thách thức kỹ thuật lớn nhất khi ứng dụng camera 2D trong điều khiển robot không chỉ là thiếu hụt thông tin chiều sâu, mà là sai số do góc lệch của camera. Trong thực tế vận hành, khi cánh tay robot tiếp cận các vật thể, chi tiết trong môi trường làm việc, trục quang học của camera hoàn toàn có thể không vuông góc với mặt phẳng nhận dạng, thao tác trên vật thể. Điều này gây ra

hiện tượng biến dạng phối cảnh (Perspective Distortion), các vật thể hình vuông, hình chữ nhật bị nhìn thành hình thang, hình tròn thành hình elip. Nếu không xác định được góc nghiêng này, robot sẽ tiếp cận vật thể sai hướng dẫn tới các sai sót khi thực thi các tác vụ trên vật thể như va chạm, kẹp lệch, kẹp trượt, khoan nghiêng ...

Trong các phương pháp xử lý ảnh hiện đại, mặc dù các mô hình học sâu (Deep Learning) như YOLO hay SSD cho kết quả nhận dạng tốt, nhưng chúng thường đòi hỏi tập dữ liệu huấn luyện lớn và hoạt động như một hộp đen, khó trích xuất chính xác các thông tin hình học 3D từ hình ảnh 2D bị biến dạng [4]. Đối với bài toán định vị vật thể theo hình dạng trong công nghiệp, các thuật toán thị giác máy truyền thống vẫn chứng tỏ được hiệu quả cao nhờ độ trễ thấp và tính ổn định. Cụ thể, thuật toán phân tích đường biên [5, 6] cung cấp cơ sở toán học vững chắc để xác định biên dạng và hướng trục chính của vật thể mà không cần phần cứng xử lý đồ họa (GPU) đắt tiền với những thuật toán phức tạp.

Xuất phát từ các phân tích trên, nghiên cứu này đề xuất một giải pháp định vị và điều khiển robot dựa trên thị giác máy trên nền kiến trúc ROS. Nghiên cứu tập trung vào hai mục tiêu là xây dựng thuật toán nhận dạng và tính toán độ lệch tư thế cho các vật thể hình học cơ bản và đồng thời thiết kế giao thức điều khiển thời gian thực để robot tiếp cận và tương tác chính xác với mục tiêu. Hệ thống

sử dụng cơ chế điều khiển vòng kín dựa trên thị giác (Visual Servoing) [7] kết hợp với kiến trúc Publisher/Subscriber [8] của ROS (Robot Operating System) để tính toán bù trừ khoảng lệch góc nghiêng theo thời gian thực đảm bảo cơ cấu chấp hành của robot luôn tiếp cận vuông góc với bề mặt cần thao tác trên vật thể. Hệ thống được kiểm chứng trên mô hình cánh tay robot 6 bậc tự do Niryo Ned 2, nhằm đánh giá tính khả thi và độ chính xác của giải pháp thị giác máy chi phí thấp trong các tác vụ thao tác thông minh.

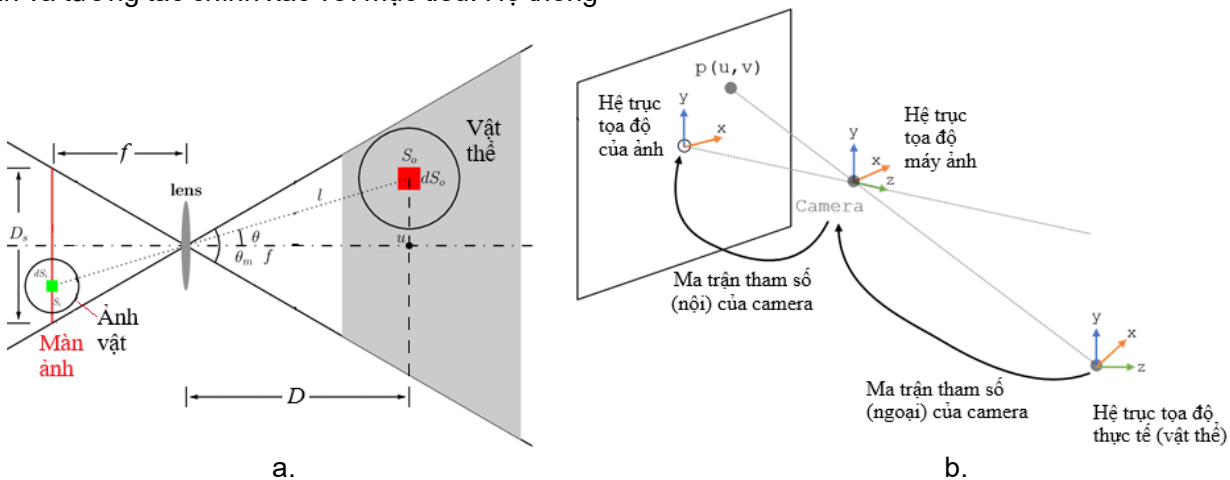
2. DỮ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Cơ sở toán học và mô hình hệ thống

Việc điều khiển robot dựa trên thị giác máy đòi hỏi sự chuyển đổi chính xác giữa các không gian tọa độ: từ không gian điểm ảnh 2D trên cảm biến camera sang không gian 3D của môi trường thực tế và cuối cùng là không gian khớp của robot.

2.1.1. Mô hình camera và ma trận camera

Để thiết lập mối quan hệ hình học giữa vật thể trong không gian 3 chiều và hình ảnh thu được, nghiên cứu sử dụng mô hình máy ảnh lỗ nhỏ - Pinhole Camera Model (Hình 1). Đây là mô hình lý tưởng mô tả cách thức hình ảnh của vật thể thực được hiển thị lên trên màn ảnh (sensor) của camera.



Hình 1. Mô hình máy ảnh lỗ nhỏ (Pinhole Camera Model) và tương quan tọa độ của vật và ảnh trong thị giác máy tính

Xét một điểm M trên vật có tọa độ (X_C, Y_C, Z_C) trong hệ quy chiếu gắn với camera (hệ trục tọa độ máy ảnh) - hình 1b, hình chiếu của nó trên mặt phẳng của cảm biến (mặt phẳng ảnh) là điểm p có tọa độ pixel (u, v) . Mối quan hệ ánh xạ trong không gian ảnh được biểu diễn qua phương trình:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} X_C \\ Y_C \\ Z_C \end{bmatrix} \quad (1)$$



trong đó: s - là hệ số tỉ lệ, tương ứng với độ sâu Z_C của điểm ảnh

K - là ma trận tham số nội (ma trận nội quan - Intrinsic Matrix) của camera, đặc trưng cho các thông số quang học và hình học bên trong của camera.

K

$$= \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

với f_x, f_y là tiêu cự tương đương của camera theo chiều cao và chiều rộng của khung hình (thường lấy xấp xỉ bằng nhau); c_x, c_y là tọa độ của tâm quang học trên màn ảnh - đặc trưng cho hệ trục tọa độ của tấm ảnh theo đơn vị pixel.

2.1.2. Chuyển đổi tọa độ

Để robot có thể tương tác với vật thể, tọa độ của vật ghi nhận được từ máy ảnh sẽ cần được chuyển từ hệ tọa độ của máy ảnh $\{C\}$ sang hệ trục tọa độ cơ sở của robot $\{O\}$. Quá trình này được thực hiện thông qua ma trận tham số ngoại của camera tương tự như mô hình sử dụng ma trận chuyển vị trong động học thuận của cánh tay robot:

$${}^O P = {}^O T_C \times {}^C P = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} \quad (3)$$

Trong đó:

${}^O P$ - là tọa độ của điểm P trong hệ trục tọa độ $\{O\}$ của robot

${}^O T_C$ - là ma trận chuyển đổi từ hệ trục tọa độ của camera $\{C\}$ về hệ trục tọa độ của robot $\{O\}$

${}^C P$ - là tọa độ của điểm P trên vật thể trong hệ trục tọa độ của camera

R - là ma trận quay, biểu diễn hướng của camera so với hệ trục tọa độ của robot

t - là tọa độ của tâm camera (tâm hệ trục tọa độ của camera $\{C\}$) trên hệ trục tọa độ của robot

Biểu diễn toán học của R và t lần lượt là các ma trận cỡ 3×3 và 3×1 :

$$(2) \quad R = \begin{bmatrix} x_{Cx} & y_{Cx} & z_{Cx} \\ x_{Cy} & y_{Cy} & z_{Cy} \\ x_{Cz} & y_{Cz} & z_{Cz} \end{bmatrix}; \quad t =$$

$$\begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} \quad (4)$$

2.1.3. Xác định khoảng cách

Trước khi xác định góc nghiêng, khoảng cách từ camera tới vật thể (D) được ước lượng dựa trên mô hình tam giác đồng dạng của thấu kính (hình 1a) - mô hình camera lỗ nhỏ (lỗ kim) - Pinhole Camera. Dựa trên kích thước thực tế đã biết của vật thể W_r và kích thước ảnh thực tế thu được W_i , khoảng cách D tương ứng với Z_c được tính theo công thức:

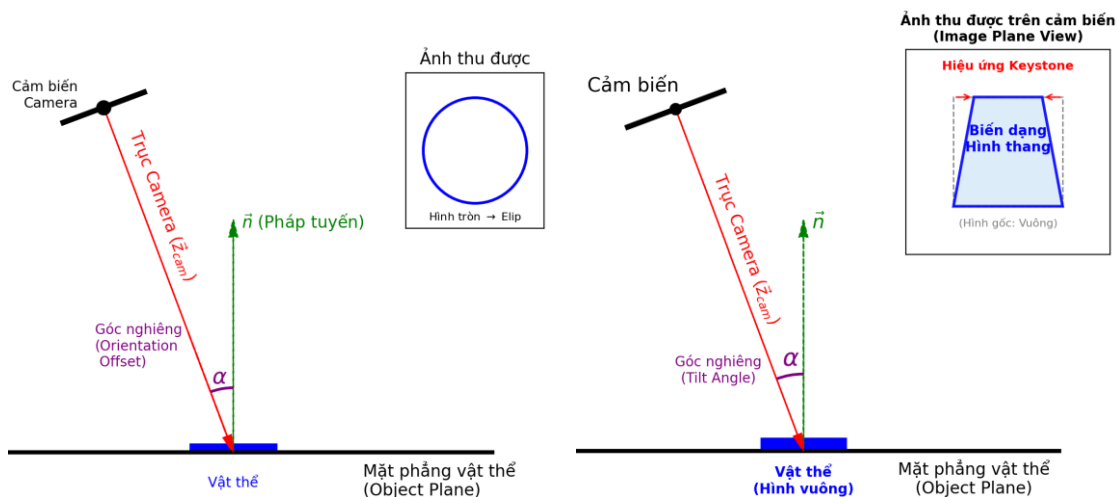
$$D = W_r \times \frac{f}{W_i}$$

$$= W_r \times \frac{f \cdot h_s}{h_a \cdot W_s} \quad (5)$$

với: h_a là độ cao ảnh theo pixel; h_s là độ phân giải theo chiều cao của cảm biến, pixel; W_s là chiều cao thực của cảm biến máy ảnh, mm.

Sai số của phép đo này phụ thuộc vào độ sắc nét của biên dạng vật thể trên ảnh.

2.1.4. Xác định góc nghiêng của camera



Hình 2. Mô hình sai số góc nghiêng α của camera gây ra biến dạng hình ảnh

Khi trục quang học của camera không vuông góc với mặt phẳng của vật thể, hình ảnh thu được

sẽ bị biến dạng như mô tả trong Hình 2: hình tròn bị biến dạng thành hình elip (Hình 1a) và hình



vuông bị biến dạng thành hình thang (hình 1b), mức độ biến dạng phụ thuộc vào độ lớn của góc nghiêng.

Để xác định góc nghiêng giữa trục camera ($\vec{z}_{cam}$) và pháp tuyến đi qua tâm của mặt vật thể ($\vec{n}$), nghiêng cứu sử dụng ma trận thuần nhất (Homography) H . Mối quan hệ giữa mặt phẳng vật thể ($Z = 0$) và mặt phẳng ảnh được mô tả bởi:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (6)$$

Ma trận H bao gồm các thành phần tịnh tiến và quay $H = K[r_1 \ r_2 \ t]$ để mô tả cách thức chuyển đổi giữa hai hệ trục tọa độ.

Vector pháp tuyến $\vec{n}$ của mặt phẳng vật thể (tương ứng trục Z của vật thể trong hệ tọa độ camera) được tính bằng tích có hướng của hai vector cột đầu tiên:

$$\vec{n} = \vec{r}_1 \times \vec{r}_2 \quad (7)$$

Góc lệch α (Tilt Angle/Orientation Offset) cần tìm chính là góc giữa trục camera $\vec{k} = [r_1 \ r_2 \ t]^T$ và vector pháp tuyến $\vec{n}$:

$$\alpha = \arccos \left(\frac{k \cdot n}{\|k\| \cdot \|n\|} \right) \quad (8)$$

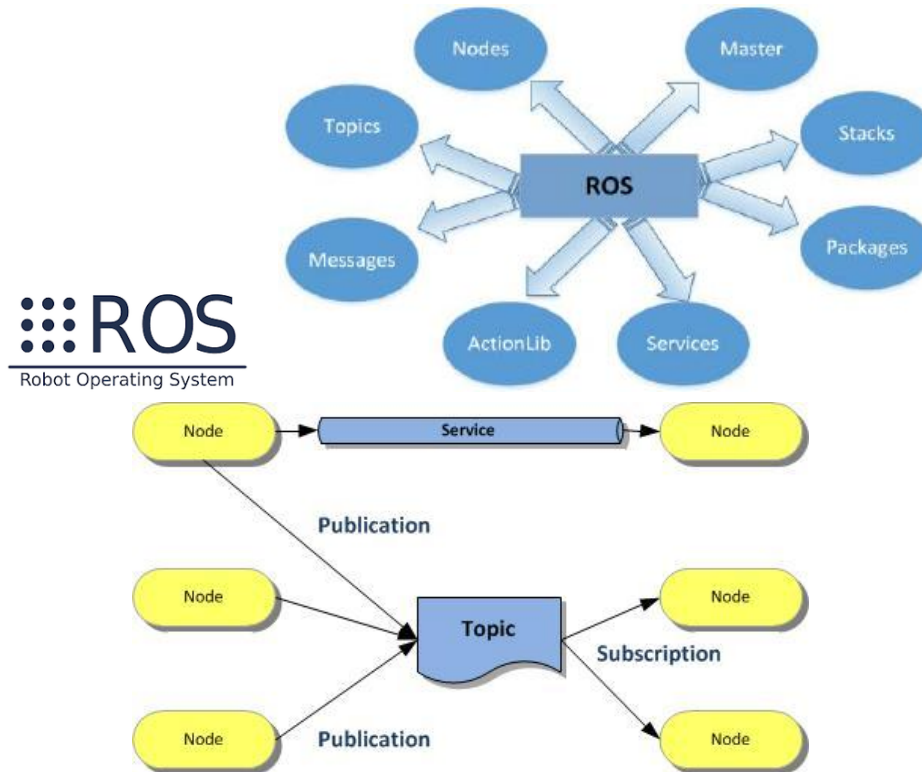
Giá trị α sẽ được gửi tới bộ điều khiển robot thông qua ROS để thực hiện các chuyển vị cần thiết theo mô hình bên dưới.

2.2. Thiết kế hệ thống điều khiển trên nền tảng ROS

Để hiện thực hóa các thuật toán tính toán khoảng lệch, góc nghiêng và điều khiển robot đã trình bày ở Mục 2, nghiêng cứu sử dụng Hệ điều hành Robot (Robot Operating System – ROS) làm nền tảng trung gian. Việc lựa chọn ROS dựa trên khả năng hỗ trợ phần cứng đa dạng và kiến trúc phân tán mạnh mẽ cho phép tách biệt hoàn toàn khâu xử lý thị giác và khâu điều khiển động học.

2.2.1. Cơ sở và kiến trúc của ROS

Hệ thống được thiết kế dựa trên mô hình của ROS, trong đó các tiến trình xử lý được chia nhỏ thành các nút (Node) độc lập, giao tiếp với nhau thông qua cơ chế gửi thông điệp (Message passing) như mô tả trong Hình 3.



Hình 3. ROS, cấu trúc và giao tiếp của các nền tảng trong ROS

Cấu trúc của hệ thống bao gồm hai phần chính:

1. Vision Node (Nút thị giác): đóng vai trò là đơn vị thị giác, nút này chịu trách nhiệm:

- Kết nối với camera để thu nhận dữ liệu hình ảnh (Raw Images)

- Thực hiện các thuật toán tiền xử lý (Canny, Thresholding)

- Tính toán ma trận đồng nhất và trích xuất các tham số hình học để xác định khoảng cách D và góc nghiêng α .



2. Control Node (Nút điều khiển): đóng vai trò là đơn vị hành động, chịu trách nhiệm:

- Nhận dữ liệu về sai số từ nút thị giác;
- Giải bài toán động học nghịch (Inverse Kinematics) cho cánh tay robot Niryo Ned 2.
- Gửi các lệnh điều khiển xuống các động cơ servo để thực hiện quỹ đạo di chuyển.

Trong nội dung của nghiên cứu này, chỉ dừng ở việc xây dựng nút thị giác, nút điều khiển chỉ xây dựng cơ chế, không thực hiện quá trình điều khiển.

2.2.2. Giao thức giao tiếp trên ROS

Nút thị giác sử dụng cơ chế Publish/Subscribe để truyền tải dữ liệu điều khiển thông qua xuất bản tin nhắn tùy biến *.msg trên topic (chủ đề). Dựa vào nội dung xử lý của nút thị giác, bản tin gửi đi sẽ bao gồm các tham số định lượng theo thời gian thực gồm: header (đồng bộ), lateral_offset (Δy), orientation_offset ($\Delta \theta$), và distance_z (Z_c). Thông thường, nút thị giác sẽ gửi (Publish) dữ liệu với tần số 30Hz.

Nút điều khiển (Control Node) thực hiện điều khiển cánh tay theo thuật toán PBVS (Position-based Visual Servoing) theo hai giai đoạn:

- Giai đoạn căn chỉnh: duy trì độ cao (khoảng cách) an toàn, thực hiện điều chỉnh vị trí và góc xoay để giảm thiểu sai số Δy và $\Delta \theta$, trong đó ưu tiên giá trị $\Delta \theta$;
- Giai đoạn tiếp cận: khi trục camera đã vuông góc với mặt phẳng vật thể, robot hạ dần khoảng

cách để tiếp cận vật thể ở đúng khoảng cách quy định.

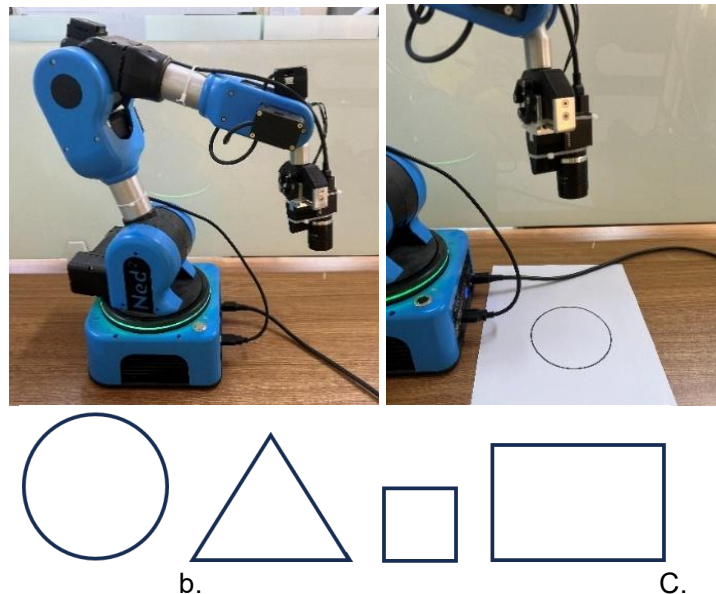
2.3. Thực nghiệm

2.3.1. Thực nghiệm

Hệ thống phần cứng bao gồm cánh tay robot 6 bậc tự do Niryo Ned 2 và camera đơn Logitech được gắn cố định trên khâu cuối theo cấu hình Eye-in-Hand (Hình 4a,b). Để đánh giá định lượng khả năng của thuật toán thị giác máy trong môi trường ROS, nghiên cứu xây dựng một tập dữ liệu chuẩn bao gồm các vật thể hình học cơ bản (hình tròn, hình tam giác, hình vuông, hình chữ nhật) được in trên giấy A4 (Hình 4c). Kích thước thực tế của các mẫu vật được chuẩn bị chính xác như sau:

- Hình tròn: đường kính $D = 100$ mm;
- Hình tam giác đều, chiều dài cạnh $a = 100$ mm;
- Hình vuông: chiều dài cạnh $a = 50$ mm;
- Hình chữ nhật: kích thước hai cạnh là 80×120 mm.

Định nghĩa điểm thao tác (Target Point): trong bài toán điều khiển tiếp cận, điểm đích mà robot cần hướng tới không phải là bề mặt của vật thể mà là một điểm nằm trên trục pháp tuyến đi qua tâm hình học, cách bề mặt vật thể một khoảng an toàn $Z_o = 200$ mm. Khoảng cách này đảm bảo cho camera có trường nhìn bao quát để tính toán xác định độ lệch, căn chỉnh hướng trước khi thực hiện các thao tác trực tiếp lên bề mặt của vật thể.



Hình 4. Mô hình thực nghiệm camera trên tay Niryo Ned 2 và tập dữ liệu vật thể mẫu

2.3.2. Quy trình xử lý ảnh trên ROS

Luồng xử lý dữ liệu tại nút thị giác được thực hiện tuần tự theo 4 bước, đảm bảo tính toán theo thời gian thực các tham số hình học:

1. Phân đoạn và trích xuất: ảnh đầu vào được xử lý để phát hiện sự hiện diện của vật thể, trích xuất đường biên (contour) và phân loại hình dạng (tròn, vuông, tam giác, chữ nhật);



2. Ước lượng khoảng cách Z_c : dựa trên kích thước thực tế đã biết của vật thể W_r , kích thước và độ phân giải của cảm biến, kích thước đo được của ảnh trên cảm biến, hệ thống tính toán được khoảng cách từ camera tới vật thể theo mô hình Pinhole camera.

3. Tính độ lệch ngang (Lateral Offset): xác định độ sai lệch Δy giữa tâm quang học của camera và tâm vật thể. Kết hợp với Z_c ta xác định được khoảng lệch thực tế giữa tâm camera và điểm thao tác.

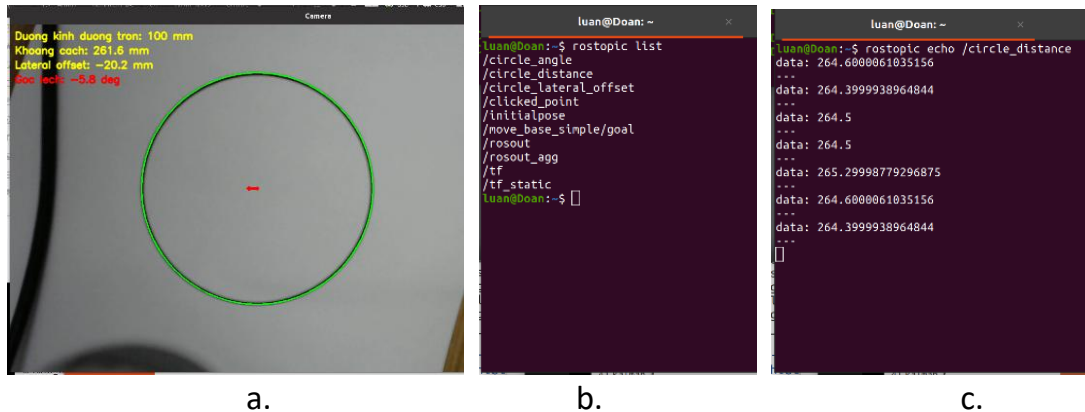
4. Tính góc lệch (Orientation/Tilt Angle): nút tính toán góc nghiêng α giữa trục quang học của camera và pháp tuyến của mặt phẳng chứa vật thể dựa trên ma trận đồng nhất (Homography).

3. KẾT QUẢ VÀ THẢO LUẬN

Các thông số tính toán như khoảng cách, khoảng lệch, góc lệch được hiển thị trực tiếp trên khung hình của camera cho phép giám sát kết quả

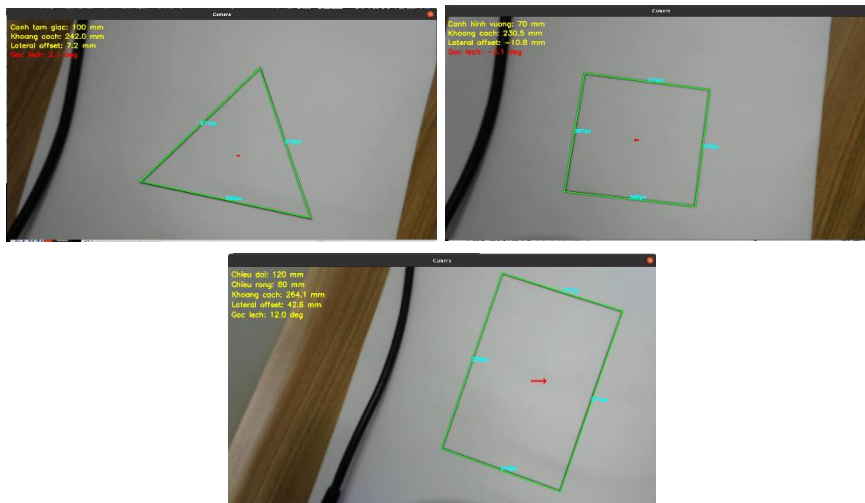
trực quan theo thời gian thực (Hình 5, Hình 6), đồng thời dữ liệu được đóng gói thành các tin nhắn để gửi lên các topic của ROS tương ứng với hình dạng vật thể và giá trị xác định được (như /circle_distance, /circle_angle ... khi nhận dạng được vật thể hình tròn – hình 5b). Dữ liệu được cập nhật liên tục với tần số cao đảm bảo độ trễ thấp cho bộ điều khiển (hình 5c).

Kết quả xử lý cụ thể đối với hình tròn: với đường kính thực 100 mm thì khoảng cách từ camera tới mặt phẳng chứa vật tính toán được là 261,6 mm, độ lệch ngang của tâm camera so với trục pháp tuyến qua tâm hình tròn là -20,2 mm với góc lệch giữa trục cam và trục pháp tuyến là -5,8°. Giá trị âm là do camera đang nằm phía dưới của trục pháp tuyến và trục quang học của camera cũng đang hướng xuống phía dưới. Kết quả cho thấy dù vật thể bị biến dạng nhẹ thành hình elip do góc nghiêng, thuật toán vẫn trả về giá trị chính xác và bám tốt với biên dạng của vật thể.



Hình 5. Kết quả xử lý hình ảnh và tính toán khoảng cách đối với vật thể hình tròn trong ROS

Khi thay đổi vị trí của cánh tay robot và chuyển vật thể cần nhận dạng sang hình tam giác, hình vuông và hình chữ nhật, kết quả tính toán được cũng rất chính xác (Hình 6), cụ thể:



Hình 6. Kết quả xử lý, tính toán với vật thể hình tam giác, hình vuông, hình chữ nhật trong ROS



Đối với hình tam giác: hệ thống nhận dạng chính xác 3 đỉnh và cạnh dài 100 mm, từ đó ước lượng được khoảng cách giữa camera với vật thể là 242 mm, sai số vị trí ngang 7,2 mm và góc lệch là $2,1^\circ$.

Đối với hình vuông: với kích thước của cạnh 50 mm, thuật toán vẫn duy trì đồ ổn định. Tại khoảng cách 230,5 mm, góc lệch tính toán được là $-3,1^\circ$ và khoảng lệch giữa tâm camera với điểm thao tác là -10,8 mm. Việc xác định đúng góc lệch này giúp robot giảm được hiệu ứng hình thang cho ảnh thu được.

Đối với hình chữ nhật: đây là trường hợp thể hiện rõ nhất hiệu quả của thuật toán xác định góc nghiêng. Kết quả thực nghiệm cho thấy góc lệch lên đến 12° và độ lệch lớn ở mức 42,6 mm. Mặc dù độ nghiêng lớn gây biến dạng đáng kể cho hình ảnh thu được, nhưng hệ thống vẫn ước lượng được kích thước cạnh 80x120 mm và khoảng cách từ camera tới mặt phẳng của vật thể là 264,1 mm một cách đáng tin cậy.

Kết quả thực nghiệm trên tập dữ liệu chuẩn cho thấy hệ thống hoạt động ổn định trong phạm vi chiều sâu (khoảng cách) từ 200 mm đến 300 mm. Sai số ước lượng khoảng cách trung bình dưới 2%. Sai số góc nghiêng phản ánh đúng thực tế gá đặt của camera, cho phép bộ điều khiển robot thực hiện bù góc chính xác. Việc tách biệt nút thị giác sử dụng giao thức ROS giúp duy trì tốc độ xử lý ảnh ổn định, không bị ảnh hưởng bởi tải tính toán của luồng điều khiển động học. Các giá trị sau tính

toán được gửi lên topic cung cấp đầu vào tin cậy cho bài toán Visual servoing.

4. KẾT LUẬN

➤ Bài báo đã đề xuất và hoàn thiện giải pháp định vị vật thể sử dụng thị giác máy đơn trên nền tảng ROS, giải quyết hiệu quả thách thức về sai số góc nghiêng trong môi trường phi cấu trúc;

➤ Kết quả nghiên cứu đạt được hai đóng góp chính: Thứ nhất, về mặt thuật toán, tác giả đã cụ thể hóa phương pháp ước lượng độ lệch tư thế dựa trên ma trận Homography. Giải pháp này cho phép xác định chính xác góc nghiêng của camera để bù trừ hiệu ứng biến dạng hình học, đảm bảo robot luôn tiếp cận vuông góc với bề mặt vật thể;

➤ Thứ hai, về mặt hệ thống, kiến trúc điều khiển phân tán theo cơ chế Publish/Subscribe đã tách biệt thành công luồng tri giác và hành động, giúp duy trì tính thời gian thực và khả năng tích hợp linh hoạt trên các dòng robot công nghiệp khác nhau mà không phụ thuộc phần cứng;

➤ Thực nghiệm trên robot Niryo Ned 2 ghi nhận sai số ước lượng vị trí trung bình dưới 2% trong vùng làm việc tối ưu và kiểm soát tốt tư thế tiếp cận. Kết quả này khẳng định tính khả thi của việc ứng dụng hệ thống thị giác 2D chi phí thấp để thay thế các cảm biến 3D đắt tiền trong các tác vụ phân loại và gắp đặt tự động hóa ngành mỏ □

TÀI LIỆU THAM KHẢO

- [1] Hashimoto, K., A review on vision-based control of robot manipulators. *Advanced Robotics*, 2003. 17(10).
- [2] Du, G., et al., Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review. *Artificial Intelligence Review*, 2021. 54(3): p. 1677-1734.
- [3] Fu, H., et al. Deep ordinal regression network for monocular depth estimation. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
- [4] O'Mahony, N., et al. Deep learning vs. traditional computer vision. in *Science and information conference*. 2019. Springer.
- [5] Suzuki, S., Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, 1985. 30(1): p. 32-46.
- [6] Hu, M.-K., Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 1962. 8(2): p. 179-187.
- [7] Chaumette, F. and S. Hutchinson, Visual servo control. I. Basic approaches. *IEEE Robotics & Automation Magazine*, 2006. 13(4): p. 82-90.
- [8] Quigley, M., et al., *ROS: an open-source Robot Operating System*. in ICRA workshop on open source software. 2009. Kobe.

LỜI CẢM ƠN

Nội dung bài báo được hoàn thành dưới sự hỗ trợ từ kinh phí từ đề tài nghiên cứu khoa học cấp cơ sở mã số T25-06 của Trường Đại học Mỏ - Địa Chất.



RESEARCH ON POSITION AND CAMERA TILT ESTIMATION FOR INDUSTRIAL ROBOT CONTROL USING MONOCULAR VISION ON ROS

Luan Cong Doan

Hanoi University of Mining and Geology, 18 Pho Vien, Ha Noi, Vietnam

ARTICLE INFOR

TYPE: Research Article

Received: 28/10/2025

Revised: 21/11/2025

Accepted: 28/11/2025

Corresponding author:

Email: doancongluan@humg.edu.vn

ABSTRACT

This paper proposes a monocular vision solution integrated with industrial robots to determine the spatial position and pose of objects within unstructured working environments. A significant challenge associated with 2D cameras is the perspective distortion error that arises when the optical axis is not perpendicular to the object's surface. To address this issue, this study develops an algorithm based on the Homography matrix to accurately estimate the Tilt Angle between the camera axis and the surface normal, while simultaneously calculating the distance and Lateral Offset using the Pinhole Camera Model. The system is implemented on the Robot Operating System (ROS) platform utilizing a distributed Publisher/Subscriber architecture, which enables the decoupling of image processing and kinematic control workflows. Experimental results on a Niryo Ned 2 robot with basic geometric objects demonstrate that the system operates stably in real-time (30Hz), achieves a distance estimation error of less than 2%, and accurately determines the tilt angle to facilitate robot approach pose compensation. This solution offers a cost-effective and highly integrable approach for automation applications in the mining industry.

Keywords: ROS, Machine Vision, Tilt Angle Estimation, Industrial Robot, Visual Servoing.

@ Vietnam Mining Science and Technology Association