VNUHCM Journal of Engineering and Technology 2026, 9(3):3079-3085

Open Access Full Text Article

Research article

A method for real-time detection of product positions on conveyors using computer vision

Thanh Huy Phung[1,2], Quang-Phuoc Tran[3,2], Phuong-Tung Pham[1,2], Quoc Chi Nguyen[1,2], Duy-Anh Nguyen[1,2]

[1]*Department of Mechatronics, Faculty of Mechanical Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet, District 10, Ho Chi Minh City 70000, Vietnam*

[2]*Vietnam National University Ho Chi Minh City (VNUHCM), Linh Trung Ward, Thu Duc, Ho Chi Minh City 70000, Vietnam*

[3]*Department of Construction Machinery and Handling Equipment, Faculty of Mechanical Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet, District 10, Ho Chi Minh City 70000, Vietnam*

History

- Received: 06-08-2023
- Revised: 09-11-2024
- Accepted: 03-07-2026
- Published Online: 04-08-2026

DOI : https://doi.org/10.32508/vnuhcmj-et.v9i3.1139

ABSTRACT

This article introduces a straightforward method for detecting the positions of products on conveyors using computer vision. The method comprises two main steps: 1) Detecting the objects and their locations by matching their distinct templates on the images, and 2) Determining the orientations via a region of interest (ROI) extracted from the vicinity of the detected locations. The objective is to reduce processing time to enable real-time applications.

For demonstration, the detection of plier pieces is used throughout the paper. A hole on the body of the plier piece is selected as the feature for detection, and the area around it is chosen as the picking location. The Hough transform is performed to detect the center of the hole, then the orientation is determined by the picking area around the hole. The proposed method performs well in both cases where a single object and multiple objects appear in the images. The method required around 40-50ms to process an individual image (1600x1200 pixels) to detect both single and multiple objects using personal computers (PCs). To increase the flexibility of the system, the method was implemented on Raspberry Pi 4. The processing time was approximately 230-240 ms in such cases.

Additionally, a method to reduce the amount of calculation and improve the process's performance is proposed. In this method, only a small region along the object's moving path is used as the region of interest (ROI) during the detection process. The location of the region is fixed, so-called the anchor location, and the detection process is only performed inside this location. In this way, the processing time is reduced to around 3-4ms with PC and 10-12ms with Raspberry Pi 4. The results are promising for real-time applications.

The proposed method can be used as a reference for applications involving other objects with complex forms in the future.

Key words: positioning, industrial, plier picking, objectdetection, pickand place

INTRODUCTION

The use of computer vision (CV) systems to automate manufacturing processes has been increasingly popular[1,2]. These vision systems serve as the "eyes" to improve the efficiency and flexibility of manufacturing processes. An essential application of CV in the industrial and manufacturing sectors is the detection of products for subsequent actions. For example, detecting the position and orientation of the products on a conveyor belt allows a robot to pick-and-place them to the desired locations.

Conventionally, a common method involves capturing images of the entire product and then calculating its centroid and orientation based on moments[3,4]. However, these processes are sensitive to noise, and the appearance of the object can be uncertain. Consequently, the detection process is typically performed in static conditions (object without motion), rather than real-time and dynamic detection (moving object).

For object detection, Deep Learning (DL) – based methods such as R-CNN and YOLO are gaining popularity[5,6] and showing promise for robotics and smart manufacturing. However, several issues need consideration. Firstly, these methods require a significant amount of computation, leading to increased processing time and cost. Additionally, data acquisition, cleaning, and model training can be time-consuming, which is not ideal for a manufacturing system. Moreover, products on production lines are often uniform and stable, presenting an opportunity to reduce computation in such cases.

Furthermore, objects with complex forms, such as non-convex shapes, may not have their centroid within their body. Consequently, even if the object and its location can be detected, defining the picking point can be challenging.

This paper proposes a simple method for object position detection to reduce calculation costs and enhance system performance. The detection is particularly useful for objects that: have complex forms, and

Cite this article : Phung T H, Tran Q, Pham P, Nguyen Q C, Nguyen D. **Ilustration of the vision system.A method for real-time detection of product positions on conveyors using computer vision**. *VNUHCM J. Eng. Technol. 2026; 9(3): 3079-3085.*

have specific shapes or features that can serve as marking templates.

Instead of detecting the whole object, the method focuses on detecting only a specific part with a distinct shape. This approach simplifies the detection process and reduces processing time.

In specific, the detection of plier pieces, which are thin and have curved non-convex forms, is presented throughout this article as an example to demonstrate the proposed method.

METHODOLOGY

Overview

Fig. 1 illustrates a typical configuration of the industrial computer vision system used in our study. The system comprises a camera equipped with a lens specially adapted for our purposes. The primary function of the camera is to detect the presence and position of objects on a moving conveyor. To ensure stable illumination during processing, front-lighting is employed[3]. The entire system is controlled by a computer as shown in Fig. 1.

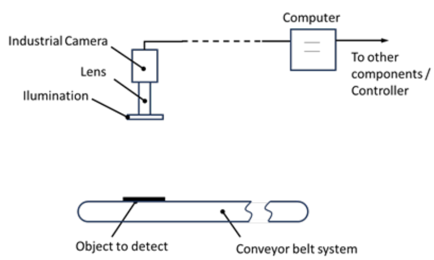

Figure 1: Ilustration of the vision system.

This paper uses the vision system to detect products on the conveyor for further actions. Fig. 3 displays the images of plier pieces, which serve as demonstration objects throughout the paper.

The conventional method for determining the object's location involves extracting the entire object from the image and then calculating its center and moments. However, this method suffers from several disadvantages. For instance, it can only detect one object at a time, limiting its efficiency in industrial settings. Moreover, the method is sensitive to noise and can only be used with static images. As shown in Fig. 2(a), (b), when an object is in motion, different parts of the object appear in the captured images at different times. This variation leads to the detection of different contours and centers for each instance, which confuses the detection process. Consequently, real-time applications are hindered.

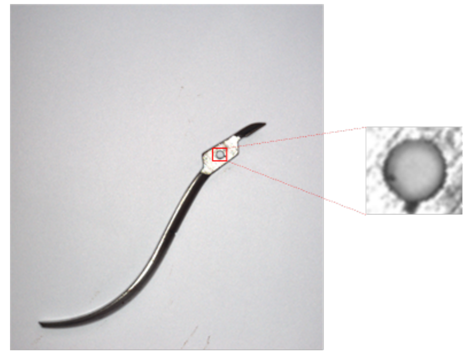

Figure 2: A typical plier pieces. (a), (b) contours of oneplier with centers (small dot) when different parts of the object appear on thecaptured image, and (c) the plier with the selected feature, which is a circlehole used for detection.

To address these limitations selecting a specific part of the object as a distinctive feature for detection, rather than analyzing the entire image. By selecting an appropriate part of the object, the presence of the object could be determined easily and reduce the influences of noise. For example, the plier piece in Fig. 3(b) has a hole in a circular form, which could be easily detected.

Detection of objects based on selected features

Fig. 3 explains the process for the detection of the objects in our method. First, the camera tried to detect the location of the selected features after system initiation. Once the location is detected, the region of interest is generated, and the orientation of the area around the feature is calculated. In this way, the convention direction could be determined and sent to the robot. Note that to determine the direction of the robot end-effector, a reference direction should be defined.

Several methods can be employed to detect the location of the features. For instance, one approach involves using images of the features to train a model specifically designed for detection. However, in our particular cases, where the features are circular, we opted to utilize the Hough transform[7] for object detection, as illustrated in Fig. 4. This choice offers the advantage of reducing sensitivity to noise in the detection process. Additionally, since the diameters of the holes vary within a known range, thresholds can be set to determine whether the detected circles correspond to the plier or are merely noise artifacts.

The direction of the object is simply calculated by the moments of the ROI extracted around the detected hole[4]. Suppose that the boundary of the object part

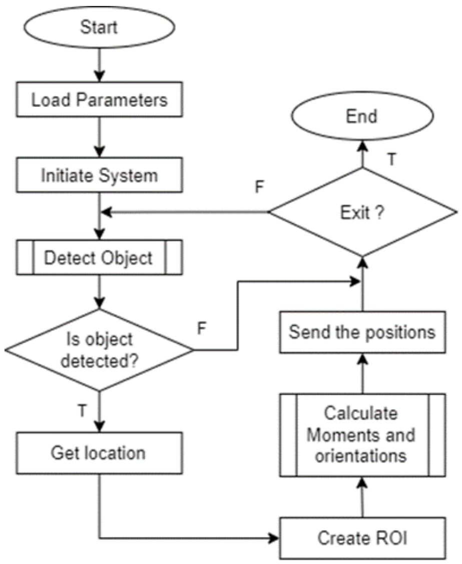

Figure 3: General detection process.

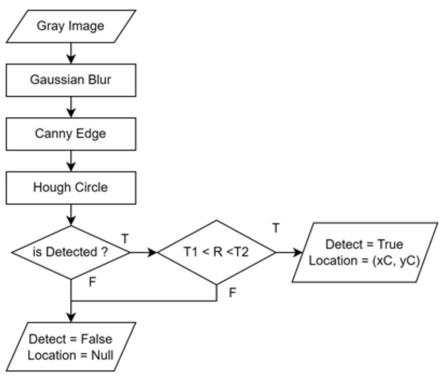

Figure 4: Object detection via selected hole.

in the ROI is a set of points $((x, y)\}$. The moment of the object can be calculated as:

$$m_{ij} = \sum_x \sum_y x^i y^j I(x,y) \tag{1}$$

Where $I(x,y)$ is the intensity of the pixel $(x,\ y)$. Because $(x,\ y)$ is on the object's boundary after edge detection, $I(x,y) = 1$ or:

$$m_{ij} = \sum_x \sum_y x^i y^j \tag{2}$$

The center coordinates of the boundary are:

$$- \tag{3}$$

Similarly, the central moment of the object part in the ROI:

$$\mu_{pq} = \sum_x \sum_y \left(x - \bar{x}_{ROI}\right)^p \left(y - \bar{y}_{ROI}\right)^q \tag{4}$$

where the indexes $p,\ q$ correspond to the order of the interested moments. The angle of the object in the ROI is determined by the angle of the principal axis:

$$\theta = \frac{1}{2} atan\left(2 * \frac{\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{5}$$

By comparing the angle with the angle of a reference, the orientation of the object could be determined.

Location anchoring

Instead of using the entire image for detection, we select only a small area across the conveyor as the Region of Interest (ROI). The width of the ROI (w)is approximately the size of the feature we want to detect, which can be represented by:

$$w = k \times d \tag{6}$$

Where the recommended coefficient $k \in (1,\ 2)$, and d is the maximum dimension of the selected feature. By focusing on this specific area, we can effectively reduce the computational burden and minimize the influence of noise or lighting variation on the detection process.

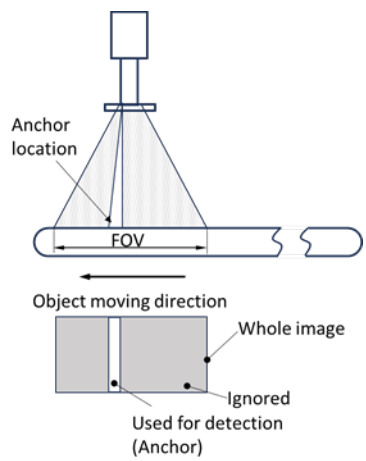

Figure 5: Anchoring method.

Experimental setup

For the demonstration of the proposed methods, images were acquired from an industrial POE (Power Over Ethernet) camera (PointGrey GS3-PGE-60S6C-C, Flir, Canada). These images have dimensions of 1600x1200 and are in 8-bit Bayer format, which can be converted from grayscale to color images. The working distance (distance between the camera lens and the objects) was about 50 cm, leading to the pixel distance (or pixel size) of about 0.15 mm. Note that the pixel distance (size) relates closely to the accuracy of the method. To ensure the pixel distance, the lens and working distance should be considered[8,9].

The computational resources utilized for the performance examination include a PC equipped with an Intel® Core^{TM} i7-4700 MQ @ 2.4GHz CPU. To ensure a fair performance comparison, we also employed Google Colab with an Intel® Xeon® @ 2.20GHz CPU. To increase the flexibility of the system, the method was also implemented on Raspberry Pi 4. All implementations were done in Python, and the computations were exclusively performed using CPUs.

Here, the examination focused solely on the algorithms and methods, without considering image acquisition and control processes.".

RESULTS AND DISCUSSION

Object detection

Fig. 6 presents the results of detection using the selected feature, specifically the hole on the body of the plier. The circle location is detected from the converted edges from the image, as illustrated in Fig. 6(a). As demonstrated in Fig. 6(a), (b), the center of the hole is easily detected within errors of ± 3 pixels, and the method exhibits noise rejection capabilities. This shows the effectiveness of the proposed approach in accurately identifying the location of the feature on the plier despite potential noise and variations in appearance.

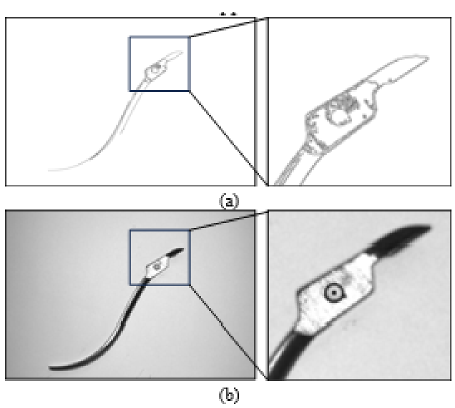

Figure 6: Circle detection by Hough Transform. (a) Converted edge for detection, and (b) detection result.

Additionally, the proposed method is versatile and capable of detecting multiple objects regardless of their intensity levels and orientations. This feature proves to be particularly beneficial in scenarios where numerous objects are simultaneously fed onto the conveyor. This further highlights the advantages and practicality of the proposed approach for real-world manufacturing and automation tasks.

Figure 7: Multiple objects detection. The objects havevaried intensity levels.

ROI extraction and Orientation calculation

After determining the location of the object through the center of the hole, a Region of Interest (ROI) can be established to calculate the orientation of the object using Eqs. (4)-(5). Fig. 8 displays the results obtained for various cases with rectangular ROIs. Note that the parameters of the ROIs should be identical for all acquired images for precise comparison, and the dimension of the ROIs should be sufficient to cover a part of the object in the vicinity of the selected feature. In this research, the ROI dimension was (170x170) pixels. In such cases, the results in Fig. 8 demonstrate that the method performs effectively under different conditions, whether for a single object or multiple objects.

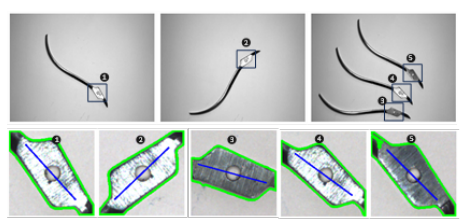

Figure 8: Calculate the oriatations of the objects via rectangular ROIs.

Considering that in the case of rectangular ROI the object portion can be affected by its location and orientation with respect to the ROI's borders, the use of circular ROI is proposed as shown in Fig. 9. The diameter of the ROI was 170 pixels for comparison with the rectangle ROI above. By employing circular ROIs, the object portion becomes orientation-invariant. This means that the shape of different acquired objects remains almost similar regardless of their orientation. As a result, the detection process using circular ROIs can lead to more stable and precise results. The variation in results between the two methods (circular ROI and rectangular ROI) is not significant (about 3 degrees). As a consequence, the rectangular ROI is still a viable option to use in certain cases. Note that in practice, a reference pose of the object should be selected, and the specific values of the direction deviation should be calculated based on the referred value:

$$\delta\theta = \theta - \theta_{ref} \tag{7}$$

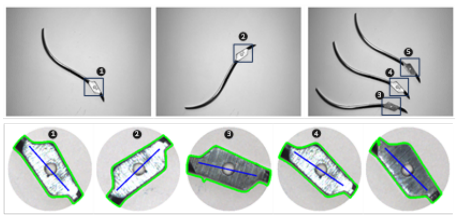

Figure 9: Calculate the oriatations of the objects via circular ROIs.

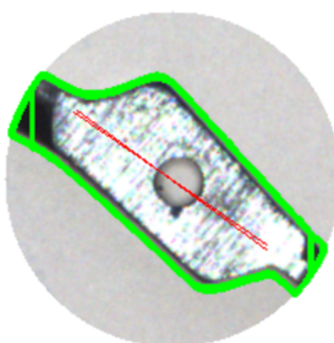

Figure 10: Slight difference between rectangle and circular ROIs.

Anchor location and performance test

To reduce processing time and stabilize object detection under various conditions such as location and illumination, the anchor location method was also explored as shown in Fig. 11. In this approach, a location shown by the camera is fixed. A region of interest (ROI) measuring 100 pixels in width was extracted from the acquired image and placed into a 1200 x 100-pixel canvas. The system exclusively detects objects within this defined area. The performance of these methods on different computer systems is presented in Fig. 12.

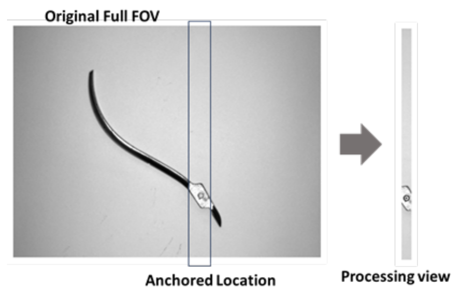

Figure 11: Slight difference between rectangle and circular ROIs.

Fig. 12 illustrates the elapsed time required for various cases and systems. As shown in Fig. 12, the PCs take about 40-50ms to process an individual image (1600x1200 pixels) for the detection of both single and multiple objects, leading to an expected maximum frame rate of 20fps. However, when using Raspberry Pi 4, the processing time increases significantly to approximately 230-240ms, yielding an expected maximum frame rate of about 4fps. As such, the performance of PCs is sufficient for real-time applications, while the Raspberry Pi falls short in meeting real-time requirements.

However, when the anchor location method is applied, the processing performance improves significantly. Even with the Raspberry Pi, the processing time is reduced to about 11-12 ms, making it suitable for real-time applications.

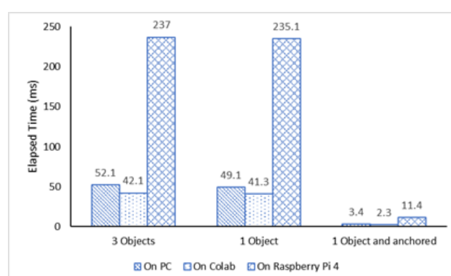

Figure 12: Elapsed time of only the algorithms.

CONCLUSION AND FUTURE WORKS

The article proposes a computer vision-based method for detecting the positions of products on conveyors, with the aim of enabling real-time applications. Additionally, the anchor location method is introduced to further improve processing time. In the demonstration with plier pieces, the method exhibited effective detection capabilities for both single and multiple objects, leveraging the Hough transform and principal axis angle. The processing time for an individual image (1600x1200 pixels) was approximately 40-50ms on PCs and 230-240ms on a Raspberry Pi 4. However, by employing the anchor location method, the processing time significantly decreased to 3-4ms and 11-12ms, respectively, making it promising for real-time applications.

Future works should involve examining the method's performance in conjunction with camera capturing and system control processes. Optimizing the method's performance by implementing it in higher-performance languages such as C++ could be beneficial. Additionally, exploring the practical accuracy of the method in specific applications, such as robot pick-and-place tasks, is essential for verification. Furthermore, experimenting with different objects and employing other detection methods, such as template matching or deep-learning-based approaches, could also be considered to enhance versatility and accuracy.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

REFERENCES

1. Mathur A, Bansal C, Chauhan S, Yadav O. A Review of Pick and Place Operation Using Computer Vision and ROS. Computational and Experimental Methods in Mechanical Engineering. 2022;p. 411–418.
2. Villalba-Diez J, Schmidt D, Gevers R, Ordieres-Meré J, Buchwitz M, Wellbrock W. Deep learning for industrial computer vision quality control in the printing industry 4.0. Sensors (Switzerland). 2019;19(18):1–23.
3. Demant C, Streicher-Abel B, Garnica C. Industrial Image Processing. Springer; 2013.
4. Bui TH, Phung T. Design and fabrication of precise shoeupper feeder system for screen printing machine in shoe industry. VNUHCM J Sci Technol Dev. 2015;18(1).
5. Glučina M, Anđelić N, Lorencin I, Car Z. Detection and Classification of Printed Circuit Boards Using YOLO Algorithm. Electronics. 2023;12(3):667–667.
6. Le PP, Nguyen VT, Guo SM, Tu CT, Lien JJJ. Visual-Guided Robot Arm Using Multi-Task Faster R-CNN. 2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI). 2019;p. 1–6.
7. Yuen HK, Princen J, Illingworth J, Kittler J. Comparative study of Hough Transform methods for circle finding. Image Vis Comput. 1990;8(1):71–77.
8. Ag B. Basler Lens Selector; 2023. Available from: https://www.baslerweb.com/en/tools/lens-selector.
9. Đ Hạnh L. Thị giác máy tính và ứng dụng trong Robotics. Nhà Xuất Bản Đại học Quốc gia; 2023.

VNUHCM Journal of Engineering and Technology 2026, 9(3): 3079-3085

Open Access Full Text Article

Bài báo nghiên cứu

Phương pháp đơn giản xác định vị trí vật thể trên băng tải theo thời gian thực sử dụng thị giác máy tính

Phùng Thanh Huy[1], Trần Quang Phước[2], Phạm Phương Tùng[1], Nguyễn Quốc Chí[1], Nguyễn Duy Ánh[1,*]

TÓM TẮT

Bài báo này giới thiệu một phương pháp đơn giản để xác định vị trí sản phẩm trên băng tải bằng thị giác máy tính. Phương pháp được đề xuất gồm hai bước chính: 1) Phát hiện các đối tượng và vị trí của chúng bằng cách so khớp các phần đặc trưng của vật thể thể hiện trên ảnh, và 2) Xác định hướng của chúng thông qua một vùng ảnh xử lý (Region Of Interest – ROI) được trích xuất từ lân cận của vị trí đã phát hiện. Mục tiêu nhằm giảm thời gian xử lý để cho phép ứng dụng thời gian thực.

Trong bài báo, việc xác định vị trí của phôi kìm được sử dụng để minh hoạ cho ý tưởng. Để ý rằng trên thân của phôi kìm có một lỗ tròn, lỗ này được chọn để xác định vị trí của phôi kìm. Phép biến đổi Hough được thực hiện để phát hiện tâm của lỗ. Sau đó, hướng của phôi được xác định bằng vùng ảnh của phôi xung quanh lỗ.

Kết quả cho thấy phương pháp đề xuất hoạt động tốt trong cả trường hợp một đối tượng và nhiều đối tượng xuất hiện trên hình ảnh. Phương pháp này cần khoảng 40-50 ms để xử lý một ảnh (1600 x 1200 điểm ảnh) xác định vị trí của một và nhiều vật thể sử dụng máy tính cá nhân. Để tăng tính linh hoạt của hệ thống, phương pháp đã được triển khai trên Raspberry Pi 4. Trong trường hợp này, thời gian xử lý vào khoảng khoảng 230-240 ms.

Để giảm lượng tính toán và cải thiện hiệu suất quy trình, bài báo đề xuất thêm phương pháp neo vị trí. Trong phương pháp này, chỉ một khu vực nhỏ dọc theo đường đi của đối tượng được sử dụng làm vùng ảnh để xử lý (ROI) và xác định vật thể. Vị trí của vùng này là cố định, gọi là vị trí neo, và quá trình phát hiện chỉ được thực hiện bên trong vùng này. Bằng cách này, thời gian xử lý giảm xuống khoảng 3-4 ms với PC và 10-12 ms với Raspberry Pi 4. Kết quả này cho thấy khả năng đáp ứng tốt của phương pháp cho các ứng dụng thời gian thực. Các phương pháp đề xuất trong bài báo này có thể được sử dụng để tham khảo khi triển khai các ứng dụng liên quan đến các đối tượng có hình dạng phức tạp khác trong tương lai.

Use your smartphone to scan this QR code and download this article

[1]*Bộ môn Cơ điện tử, Khoa Cơ khí, Trường Đại học Bách Khoa, Đại học Quốc gia Thành phố Hồ Chí Minh*

[2]*Bộ môn Kỹ thuật Máy xây dựng và Nâng chuyển, Khoa Cơ Khí, Trường Đại học Bách Khoa, Đại học Quốc Gia Thành phố Hồ Chí Minh*

Liên hệ

Nguyễn Duy Ánh, Bộ môn Cơ điện tử, Khoa Cơ khí, Trường Đại học Bách Khoa, Đại học Quốc gia Thành phố Hồ Chí Minh

Email: ndanhcdt@hcmut.edu.vn

Lịch sử

- Ngày nhận: 06-08-2023
- Ngày sửa đổi: 09-11-2024
- Ngày chấp nhận: 03-07-2026
- Ngày đăng: 04-08-2026

DOI : https://doi.org/10.32508/vnuhcmj-et.v9i3.1139

Trích dẫn bài báo này: Huy P T, Phước T Q, Tùng P P, Chí N Q, Ánh N D. **Phương pháp đơn giản xác định vị trí vật thể trên băng tải theo thời gian thực sử dụng thị giác máy tính**. *VNUHCM J. Eng. Technol. 2026; 9(3): 3079-3085.*