VISION-BASED 3D OBJECT COORDINATE ESTIMATION AND TRACKING USING HOUGH TRANSFORM AND MODULAR 3D ESTIMATION IN ROS

Oussama Errouji[1], NACIRI Imad-Eddine[1], and Jade Bousliman[1]

[1]Euromed University of Fez

October 09, 2025

VISION-BASED 3D OBJECT COORDINATE ESTIMATION AND TRACKING USING HOUGH TRANSFORM AND MODULAR 3D ESTIMATION IN ROS

A PREPRINT

Oussama Errouji
Euromed University of Fez

NAICRI Imad-Eddine
Euromed University of Fez

Jade Bousliman
Euromed University of Fez

September 22, 2025

ABSTRACT

This paper presents a ROS2-based robotic system for real-time ball detection, 3D localization, and autonomous tracking. Departing from traditional blob detection approaches, the system employs the Hough Circle Transform for robust 2D detection combined with a decoupled architecture where 3D estimation operates independently from motion control. The modular design enables simultaneous applications: reactive tracking via 2D data and 3D spatial analysis for external tasks. Extensive evaluation in Gazebo simulations demonstrates system adaptability to varying lighting conditions, ball sizes, and distances. Potential applications span automated ball collection, industrial sorting, and augmented reality interactions.

Keywords ROS2 · Hough Transform · Monocular Depth Estimation · Gazebo Simulation · Robotic Vision

1 Introduction

Robust detection and tracking of spherical objects in dynamic environments remains a critical challenge in robotics. This work advances prior methodologies through three principal contributions:

1. Implementation of the **Hough Transform** as a powerful and efficient algorithm for accurate circle detection under noisy conditions
2. Decoupled architecture enabling independent 3D estimation and tracking processes for parallel processing flexibility
3. Comprehensive validation through systematic Gazebo benchmarks including multi-ball tracking and varied lighting conditions

The system's ROS2-based architecture ensures scalability, while simulation-driven evaluation guarantees reproducibility and robustness across diverse operational scenarios.

2 Methodology

2.1 System Architecture

The proposed pipeline comprises three independent nodes (Fig. 1):

- `detect_ball`: 2D detection using Hough Transform
- `detect_ball_3d`: 3D position estimation
- `follow_ball`: Reactive velocity control

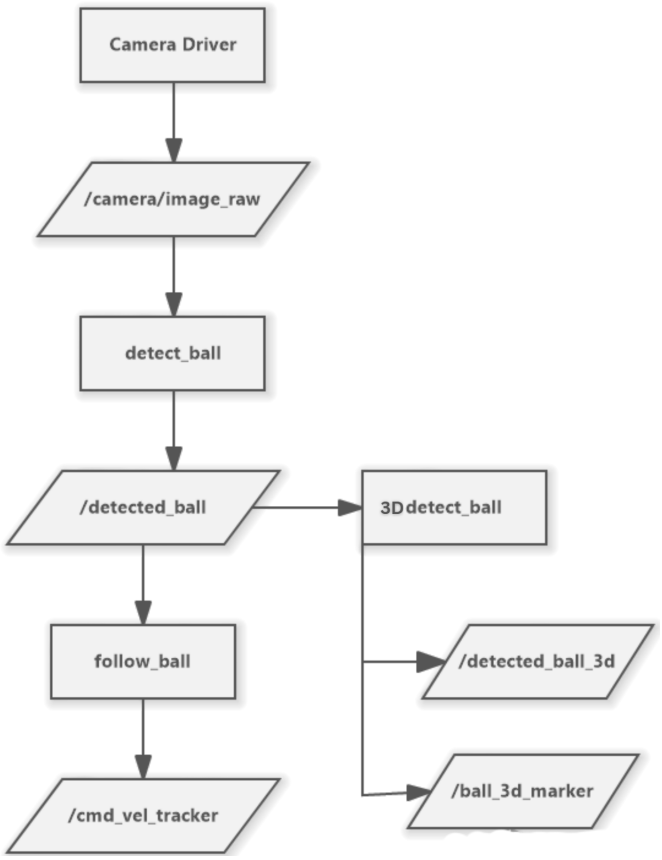

Figure 1: System architecture demonstrating modular node communication

2.2 2D Detection via Hough Transform

The `detect_ball` node processes images from `/camera/image_raw` through the following pipeline:

- **Preprocessing**: Gaussian blur application for noise reduction, followed by Adaptive Histogram Equalization to enhance local contrast and edge visibility
- **Color Masking (HSV)**: Image conversion to HSV color space with thresholding to isolate target hue ranges, ensuring robustness to lighting variations
- **Hough Circle Transform**: Circle detection using gradient-based Hough Transform, which maps edge points into a three-dimensional accumulator space (a, b, r) corresponding to potential circle centers (a, b) and radii r. A circle with center (a, b) and radius r satisfies:

  $$(x - a)^2 + (y - b)^2 = r^2$$

  Edge pixels vote in this parameter space, with peaks in the accumulator indicating likely circle candidates. Filtering is applied using:
  - *Minimum/maximum radius* constraints to discard irrelevant detections
  - *Accumulator threshold* to select high-confidence candidates
- **Output**: Publication of normalized pixel coordinates (u, v) and detected radius to `/detected_ball`

2.3 3D Position Estimation

The `detect_ball_3d` node subscribes to `/detected_ball` and calculates 3D coordinates (x_{3d}, y_{3d}, z_{3d}) using geometric relationships:

Depth Estimation from Apparent Size

$$\theta_{\text{ball}} = z_{2d} \times \text{hfov} \tag{1}$$

$$d = \frac{r}{\tan\left(\frac{\theta_{\text{ball}}}{2}\right)} \tag{2}$$

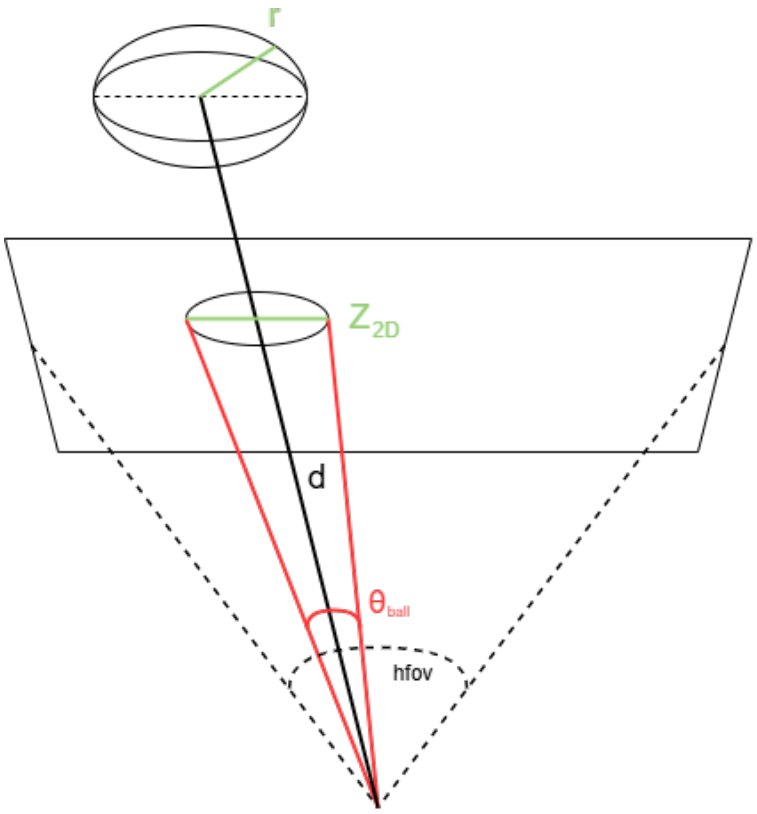

Figure 2: Geometric relationship for depth estimation using apparent size and horizontal field of view

Vertical Angle and Y Coordinate Calculation

$$\theta_y = y_{2d} \times \frac{\text{vfov}}{2} \tag{3}$$

$$y_{3d} = d \times \sin(\theta_y) \tag{4}$$

$$d' = d \times \cos(\theta_y) \tag{5}$$

Horizontal Angle and X/Z Coordinate Calculation

$$\theta_x = x_{2d} \times \frac{\text{hfov}}{2} \tag{6}$$

$$x_{3d} = d' \times \sin(\theta_x) \tag{7}$$

$$z_{3d} = d' \times \cos(\theta_x) \tag{8}$$

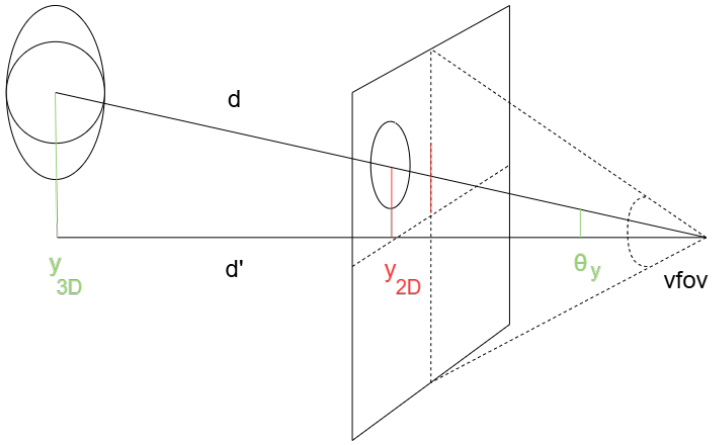

Figure 3: Vertical position estimation and projected distance calculation

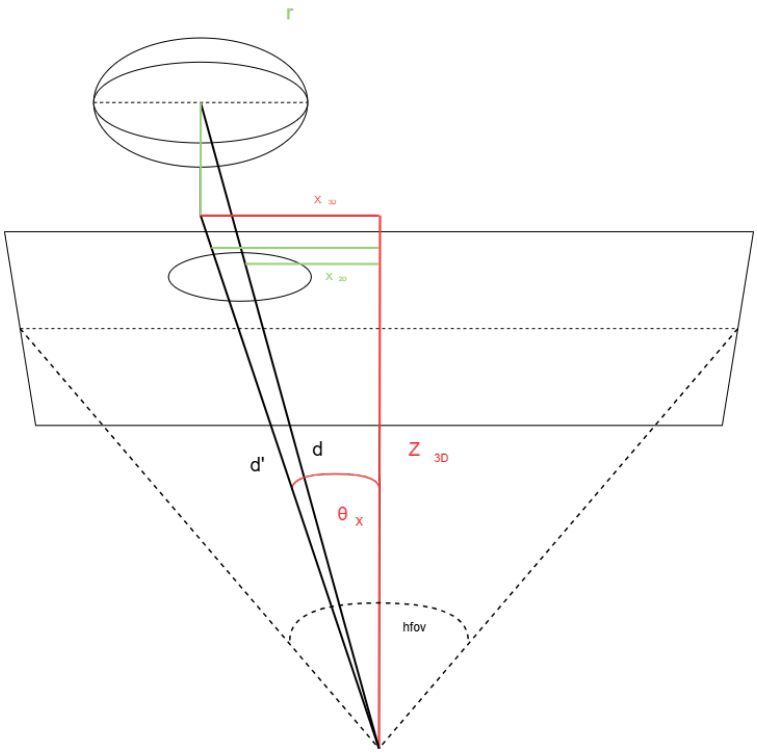

Figure 4: Horizontal position estimation using angular relationships

2.4 Reactive Tracking Control

The tracking system employs a reactive closed-loop control strategy that continuously adjusts robot motion based on real-time visual feedback:

- **Angular Control**: Proportional (P) controller minimizes horizontal displacement between ball position and camera center. For normalized horizontal offset $x \in [-1, 1]$ ($x = 0$ indicates perfect alignment), angular velocity ω is computed as:

$$\omega = -K_p \cdot x$$

  where K_p represents the angular gain (chase multiplier)

- **Linear Control**: Forward velocity v is regulated based on perceived distance to the ball, using pixel radius r as an inverse distance proxy. Forward motion continues only when:

$$r < r_{\text{max}}$$

with constant velocity $v = v_f$ until reaching desired proximity

- **Search Behavior**: Upon detection timeout $\Delta t > t_{\text{max}}$, the robot enters search mode with fixed angular velocity ω_s for target reacquisition
- **Exponential Filtering**: Smoothing of horizontal offset x and radius r values through:

$$\hat{x}_t = \alpha \cdot \hat{x}_{t-1} + (1 - \alpha) \cdot x_t$$

where $\alpha \in [0, 1]$ controls filtering memory

3 Results

3.1 Detection Robustness Under Varying Illumination

System reliability was assessed under varying light intensities by adjusting diffuse RGB values of the scene's main light source. Approximate lux levels were computed using:

$$\text{Lux} \approx \left(\frac{R + G + B}{3} \right) \times L_{\text{ref}} \tag{9}$$

where $R, G, B \in [0, 1]$ represent normalized diffuse color components and $L_{\text{ref}} = 10{,}000$ lux for standard white light.

Detection results (Table 1) demonstrate reliable operation down to approximately 500 lux, confirming robustness to ambient lighting variations.

Diffuse RGB	Approx. Lux Level	Detected
`RGB(0.8, 0.8, 0.8)`	∼8000 lx	True
`RGB(0.5, 0.5, 0.5)`	∼5000 lx	True
`RGB(0.1, 0.1, 0.1)`	∼1000 lx	True
`RGB(0.05, 0.05, 0.05)`	∼500 lx	True

Table 1: Detection performance across illumination conditions

3.2 Spatial Detection Range

Operational range testing established:

- **Minimum distance:** 10 cm (limited by field-of-view and distortion constraints)
- **Maximum distance:** 3.5 m (limited by camera resolution and circle size threshold)

This range proves adequate for indoor robotic applications including following, pick-and-place, and reactive interaction tasks.

3.3 Coordinate Output and Tracking Performance

Figure 5 shows sample detection output with 3D coordinates:

- $x = 0.15$ m
- $y = -0.34$ m
- $z = 0.045$ m

3.4 Velocity Control Response

Figure 6 demonstrates motion behavior during tracking, showing initial rotational alignment followed by steady forward approach.

Figure 7 illustrates successful approach and alignment between initial time t_0 and final time t_1.

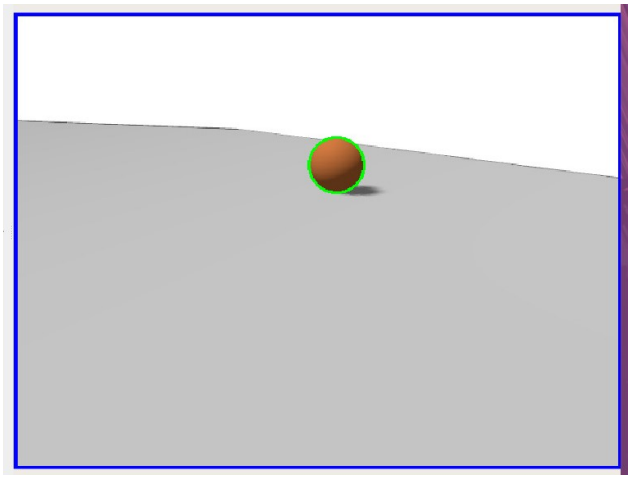

Figure 5: Ball detection with overlayed coordinate output

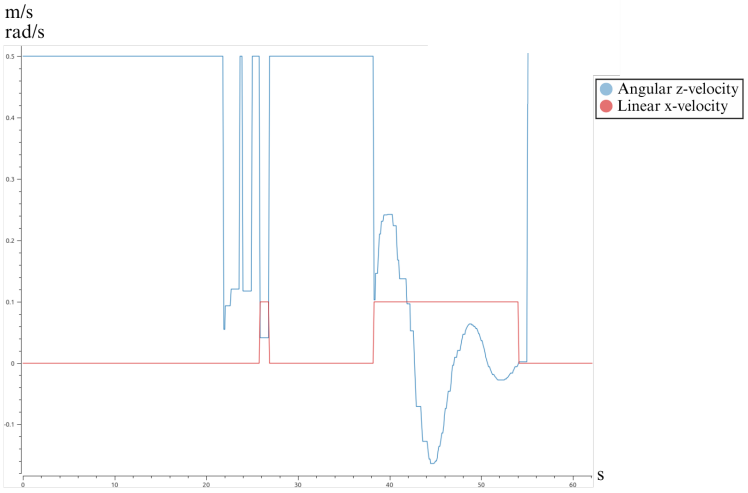

Figure 6: Linear and angular velocity profiles during tracking operation

3.5 Trajectory Analysis

Temporal evolution of ball coordinates (camera frame) and robot trajectory (world frame) in Figure 8 confirms smooth target approach with consistent distance reduction and minimal lateral drift.

4 Applications

The developed system enables numerous practical applications:

- **Automated Ball Collection**: Robotic gathering of sports balls in stadium environments using 2D tracking
- **Industrial Sorting**: 3D coordinate guidance for robotic arms to categorize objects by size and spatial position
- **Augmented Reality**: 3D data integration for overlaying virtual elements on real-world ball games
- **Agricultural Robotics**: Size-filtered Hough detection for fruit-picking robots in harvesting operations

5 Conclusion

This work demonstrates a modular ROS2 framework for vision-based ball tracking, validated through rigorous simulation benchmarks. The decoupled architecture supporting simultaneous 2D and 3D processing enables diverse ap-

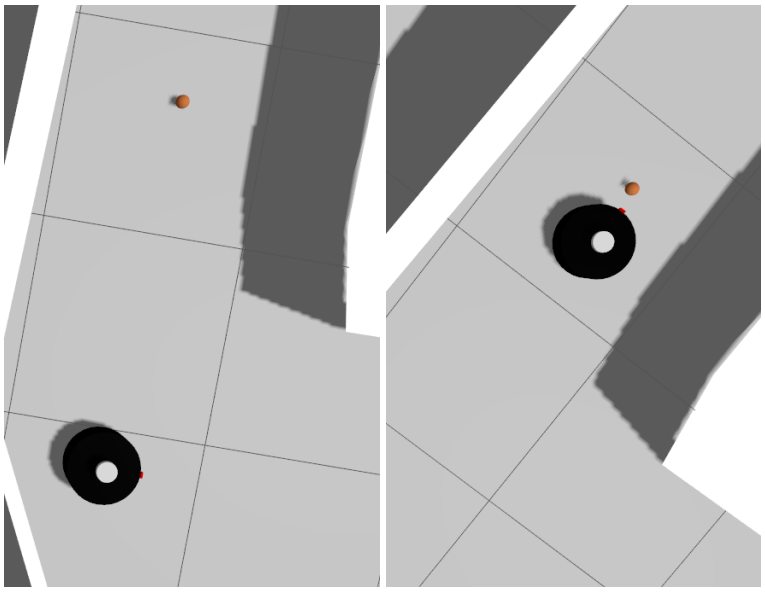

Figure 7: Robot position at t_0 (left) and t_1 (right) demonstrating successful tracking

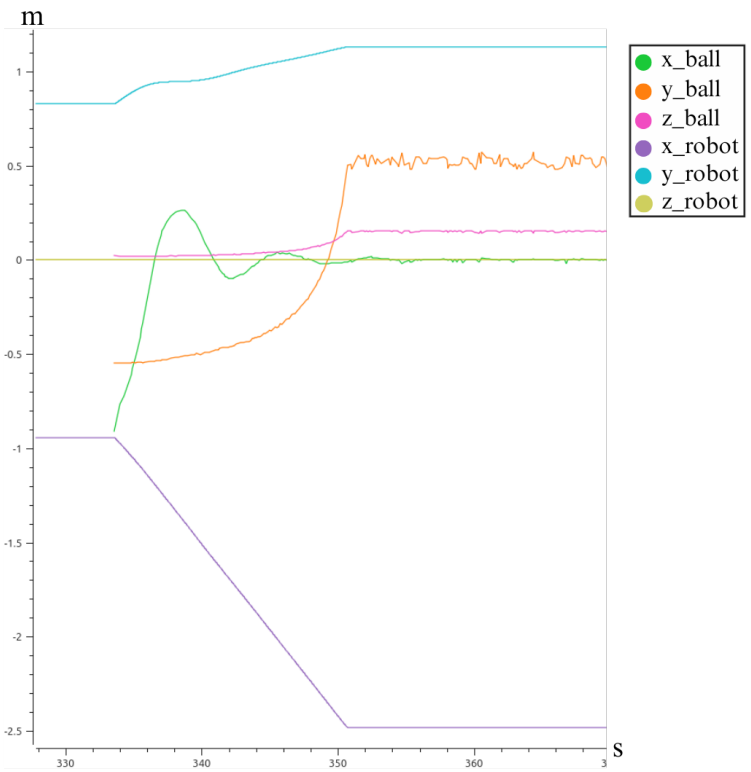

Figure 8: Temporal evolution of coordinates and trajectory

plications ranging from industrial automation to interactive augmented reality. Future work will focus on real-world validation and multi-object tracking capabilities.

References

[1] Ioan Ungurean, "Timing Comparison of the Real-Time Operating Systems for Small Microcontrollers," *Symmetry*, vol. 12, no. 1, pp. 56, 2020.

[2] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000. [Online]. Available: `https://docs.opencv.org`

[3] D. Eigen, C. Puhrsch, and R. Fergus, "Depth Map Prediction from a Single Image using a Multi-Scale Deep Network," *NeurIPS*, 2014.

[4] D. Thomas, W. Woodall, and B. Biggs, "Next-generation ROS: Why ROS 2?," *Open Source Robotics Foundation*, 2018. [Online]. Available: `https://design.ros2.org/`

[5] N. Koenig and A. Howard, "Design and Use Paradigms for Gazebo, An Open-Source Multi-Robot Simulator," *IEEE/RSJ IROS*, 2004.

[6] P.V.C. Hough, "Method and Means for Recognizing Complex Patterns," U.S. Patent 3,069,654, 1962.

[7] R. Kelly, V. Santibañez, and A. Loría, *Control of Robot Manipulators in Joint Space*, Springer, 2005.

[8] A. Yilmaz, O. Javed, and M. Shah, "Object Tracking: A Survey," *ACM Computing Surveys*, vol. 38, no. 4, 2006.

[9] J. T. Kent, "Exponential Filtering of Time Series Data," *Statistical Science*, vol. 1, no. 3, 1986.

[10] M. Billinghurst, A. Clark, and G. Lee, "A Survey of Augmented Reality," *Foundations and Trends in HCI*, vol. 8, no. 2-3, 2015.

[11] T. Bac et al., "Harvesting Robots for High-value Crops: State-of-the-art Review and Challenges Ahead," *Journal of Field Robotics*, vol. 31, no. 6, 2014.

[12] A. Hernandez, S. Cousins, "Real-time Control in ROS: Enabling Multi-Robot Interaction," *ROSCon*, 2013.

[13] M. Bhowmik, A. Kundu, and B. Bhowmik, "Low-Light Image Enhancement for Object Detection Using CNN," *Procedia Computer Science*, vol. 171, 2020.