3D POINT CLOUD-BASED PALLET DETECTION AND POSE ESTIMATION FOR INDUSTRIAL MOBILE ROBOTS

Samantha Caballero Gómez

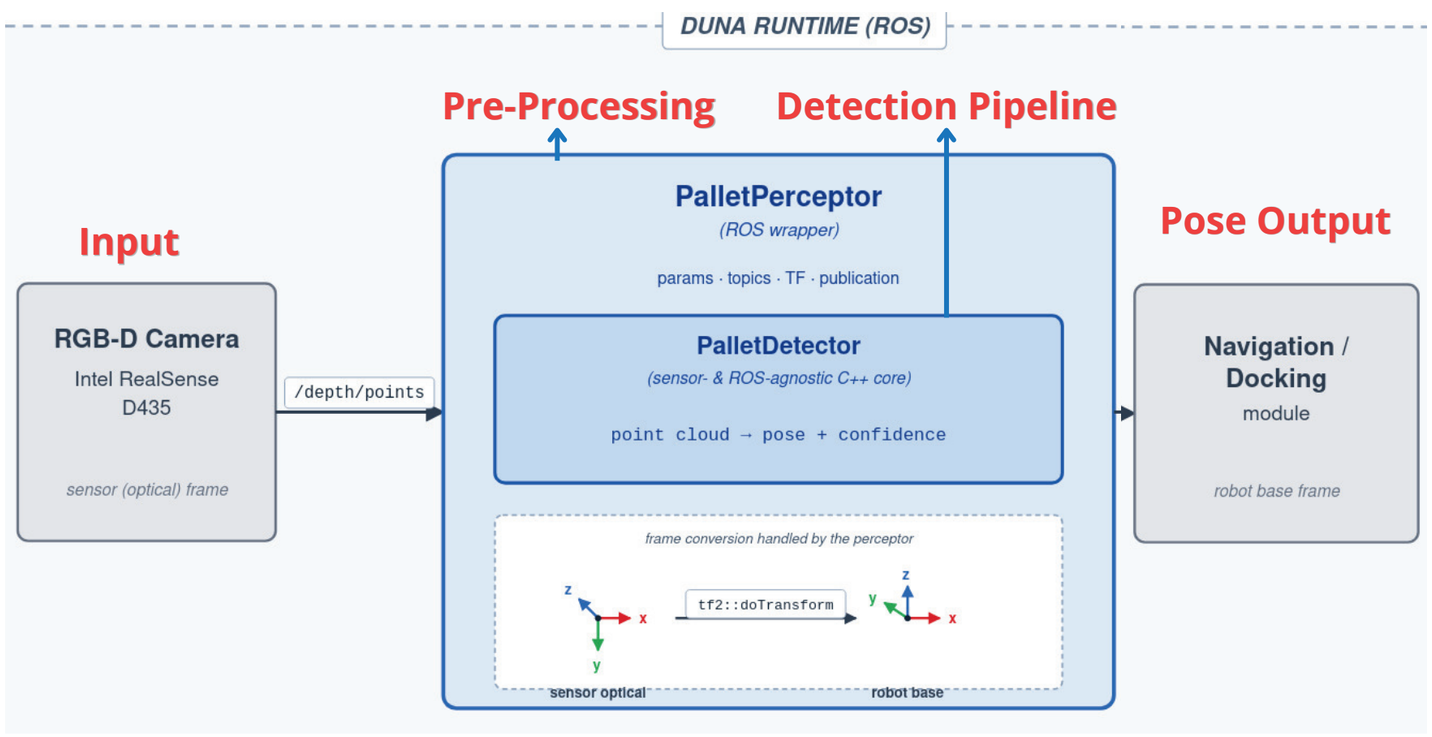
Supervisors: PhD Andreu Corominas | Prof. Levente Hajder

Universitat de Girona · BETA ROBOTS · IFRoS Erasmus Mundus Master in INTELLIGENT FIELD ROBOTIC SYSTEMS · ELTE EÖTVÖS LORÁND UNIVERSITY

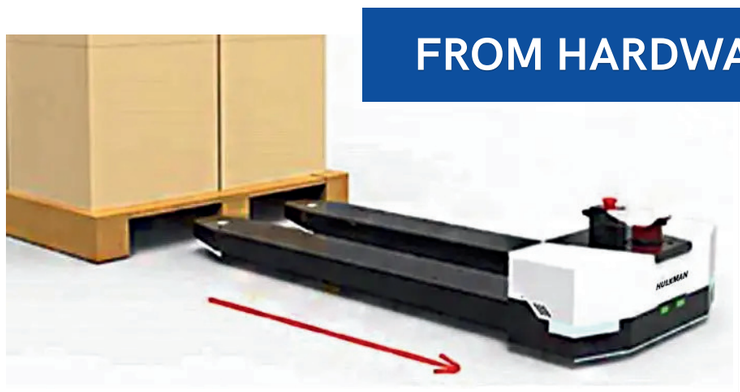
OVERVIEW

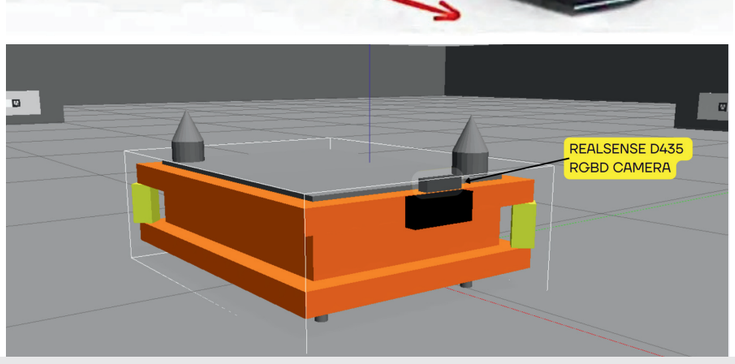
Industrial mobile robots pick pallets that rarely sit at fixed coordinates, so the robot must first recover the pallet's full pose from its onboard camera. This thesis takes a purely geometric approach to estimate the pallet pose from a single RGB-D point cloud, using Euclidean clustering, pallet-face template matching, and RANSAC plane fitting. The output is a pose that downstream navigation can use to drive the robot to the pallet. The pipeline is implemented as a PCL-based ROS node and validated in Gazebo simulation. The system is split into a sensor- and ROS-agnostic detection core that handles only the geometric algorithm, and a ROS wrapper that handles parameter loading, frame transforms, and publication, keeping the algorithm reusable and the integration code isolated.

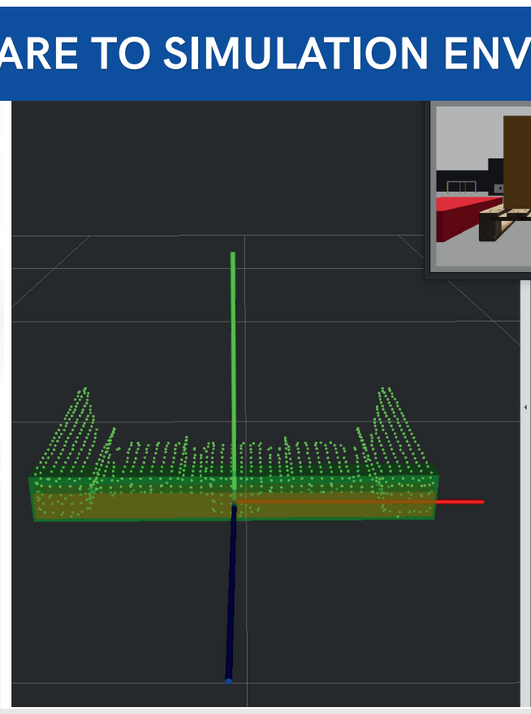
SYSTEM ARCHITECTURE

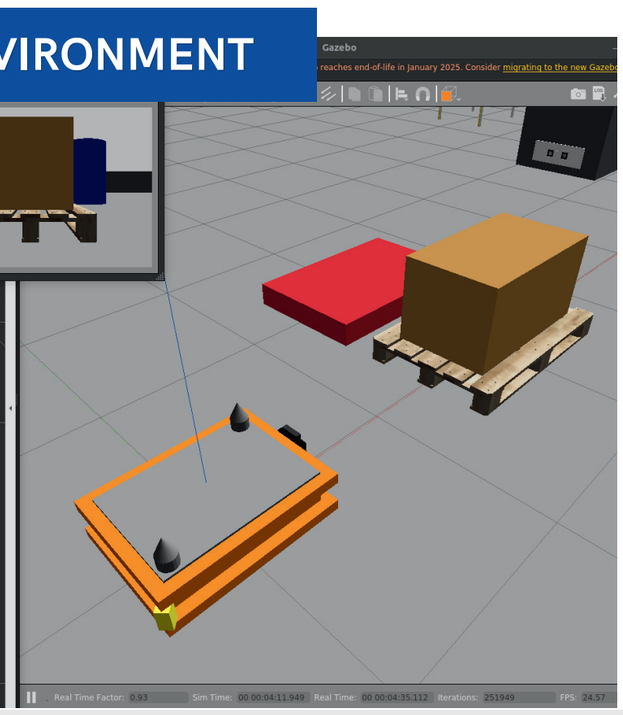


FROM HARDWARE TO SIMULATION ENVIRONMENT









The detector is tested in Gazebo simulation that reproduces the D435 camera, the EUR/EPAL pallet, and a warehouse-style scene with clutter and decoys.
Because the simulator's built-in depth noise is unrealistically clean, a custom noise-injector node perturbs each point with a range-dependent model that matches how a real D435's depth error grows with distance.

PALLET DETECTION PIPELINE

Cloud preprocessing

front end · runs once per frame

1. **Y-band crop** — restrict to pallet height band
2. **Voxel downsample** — 2 cm leaf, uniform density
3. **Euclidean clustering** — segment candidate objects

output: candidate clusters

clusters

Per-cluster recognition

loop body · runs per candidate cluster

A. **Statistical outlier removal** — clean the cluster before fitting

B. **Face-normal RANSAC** — find pallet face direction · ±25° gate

C. **EUR/EPAL template match** — project to 2D grid, slide binary template, score χ

D. **Width gate** — reject if matched width ≠ 0.80 ± 0.10 m

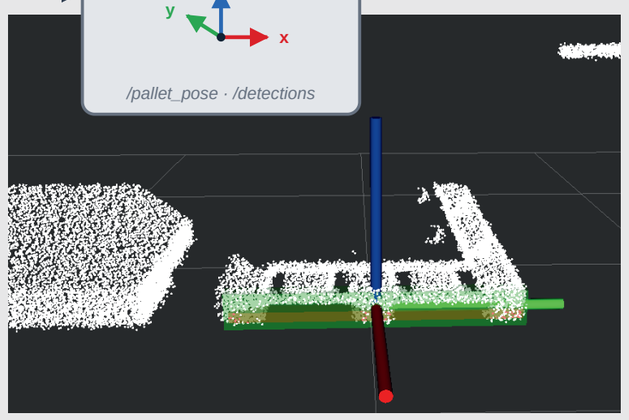
E. **Pose-RANSAC + yaw check** — stringer-zone fit → yaw, x, z · consistency check

best χ

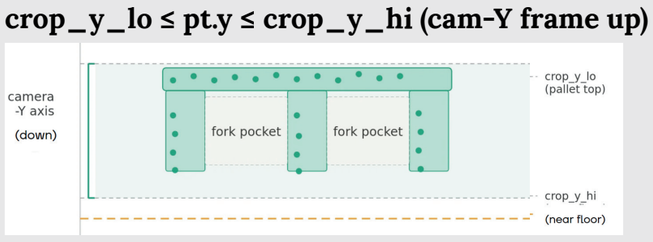
Published pose

at front-face centre

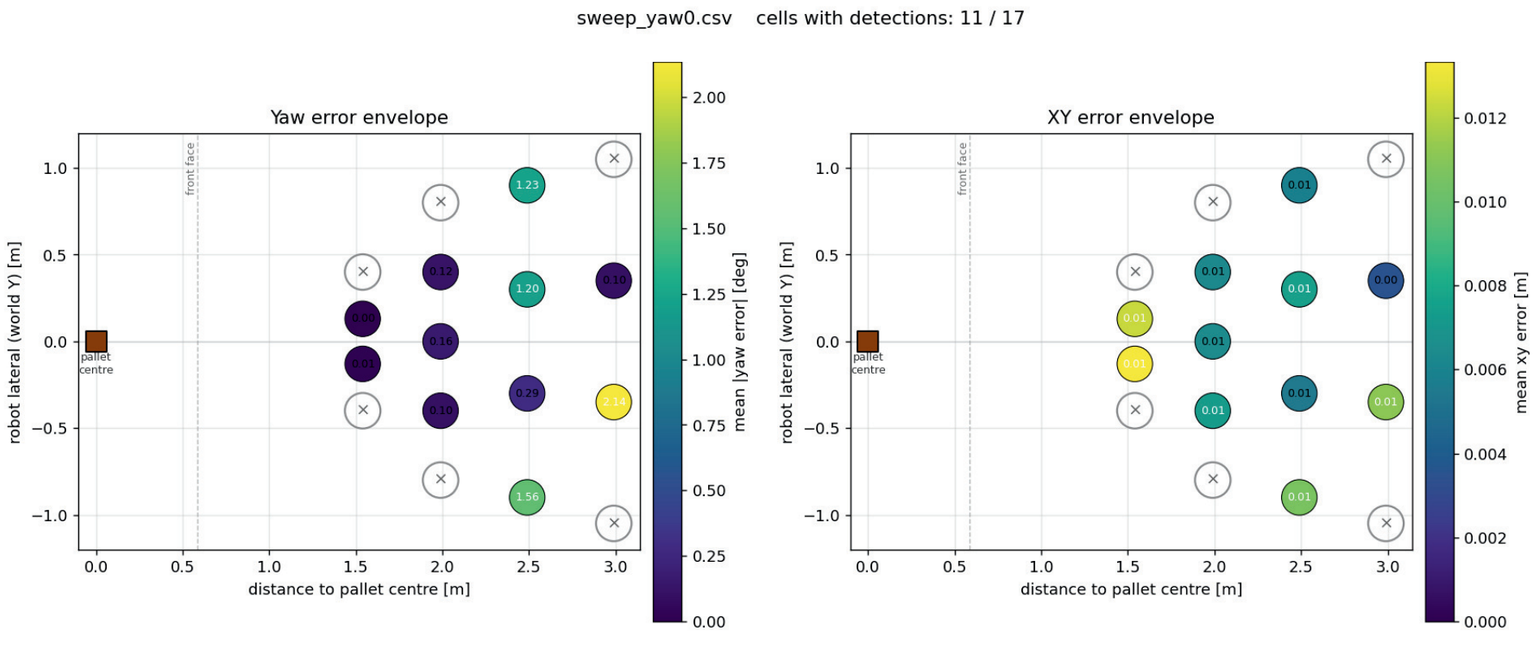
/pallet_pose · /detections



Height crop → Keeps face points where crop_y_lo ≤ pt.y ≤ crop_y_hi (cam-Y frame up)



Running once per point cloud, the scene is first cropped to the pallet's height band, downsampled, and split into candidate clusters. The pallet is identified by matching its front-face shape against a binary EUR/EPAL template, and its orientation refined by fitting a plane to the matched points (RANSAC).



RESULTS AND CONCLUSION

The detector's operating envelope was mapped by sweeps that teleport the robot across a grid of positions at fixed yaws, to get per-cell translation and yaw error vs. ground truth.
The detector holds centimetre-level planar and sub-degree yaw accuracy on fresh detections, with failures concentrated only where the pallet face leaves the camera's field of view, confirming reliable pose estimation throughout the envelope it is designed to operate in (-20° to +20° of yaw at roughly 1.5-2m robot-to-pallet euclidean distance).