Compsys 732 Project 1 report

Introduction

Project 1 aims to determine which features could be used as landmarks for the localisation and mapping of the robots. In the world of robotics, a robot has to understand and navigate its environment and the efficiency and accuracy of this task depend on how well the robot can detect its surroundings. Among various sensors, depth-sensing cameras have been used for this project which can not only capture the visual appearance but also get the distance to objects. This provides the robot with an understanding of the 3D space around them. Landmarks play a vital role in localisation and mapping. This project aims to determine which features in the lab/testing environment can be used as landmarks for the robot to utilise for localisation and mapping purposes. For this project, I evaluated three potential landmarks: cones, boxes and wall corners as they are static reference points. The testing environment is set up as a gazebo virtual simulation and the world with some cones and boxes was used.

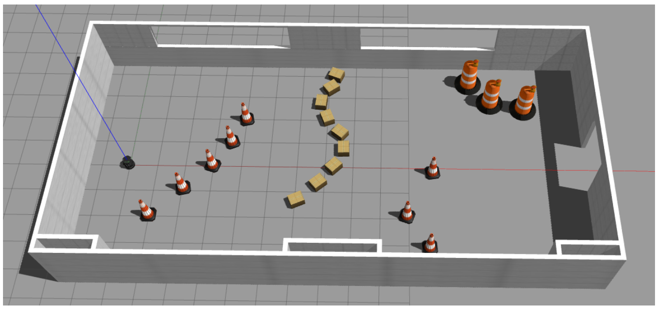

Fig.1. Gazebo virtual environment used for this project

Assumptions

The cones have a cylindrical form with a pointed tip. They are typically made of plastic and they are in two different shapes in the tested environment. As they have distinct shapes, they will be easily and clearly identified as in depth images. However, as they are not anchored, the location of them could be changed by the collision with the robot. Assuming that the robot does not push the cone, due to its shape, I assume that it is appropriate to be used as one of the landmarks. The boxes used in the testing environment are made of cardboard and in rectangular prisms. They are easy to detect as they have sharp edges and present clear shapes in depth images. However, similar to cones, as they are not anchored, their position could possibly change by the collision with the robot. Additionally, I assume the wall corners give consistent and unique features to be used as landmarks.

Implementation

For the project, the depth camera is being used, the topic "/camera/depth/image_raw" is subscribed and the message type used is sensor_msgs/Image as shown in Fig.7. The node subscribes to the depth image from the camera. This image represents the distances between the camera and the objects in its field of view, with each pixel's value corresponding to a distance.

For each depth image, it is converted to OpenCV format and then normalised for visualisation purposes as mentioned in. The binary images are created for contour detection and their properties such as centroid, and area are calculated. The 'find_depth_point' function is used to compute the 3D position of the centroid of each contour. The 'find_depth_point' function converts a 2D pixel coordinate in an image to a 3D point in space. It uses the camera's intrinsic parameters (fx, fy, cx, cy) and the depth value (z) at the pixel's location to compute the corresponding 3D coordinates (x,y,z). The result is formatted as a dictionary representing a 3D pose with position and orientation.

The corners are also detected using the 'detect_corners' function. It detects corners in a binary image and returns the detected corner coordinates. The distance is also being measured by using the depth value of the pixel. It represents the distance from the camera to the object at the specific pixel location. For each detected contour, the depth value at its centroid is extracted and displayed on the processed image. The depth values at the detected corners are also displayed. The ‘find_depth_point’ function also converts this depth value (in combination with pixel coordinates and camera intrinsic parameters) to a 3D point in space.

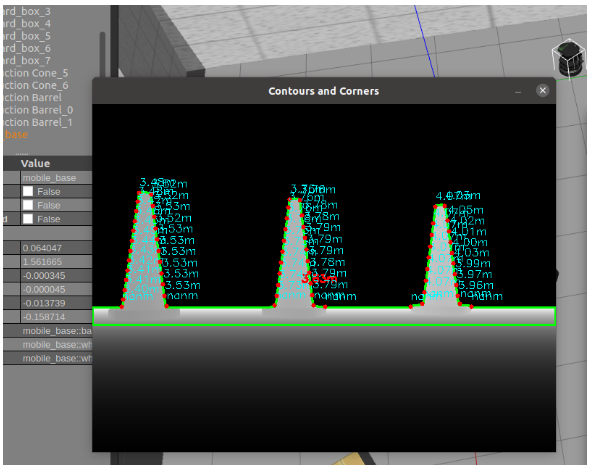

Fig.2. Small cones

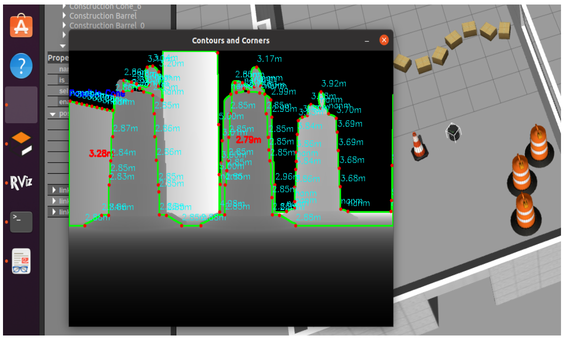

Fig.3. Big cones

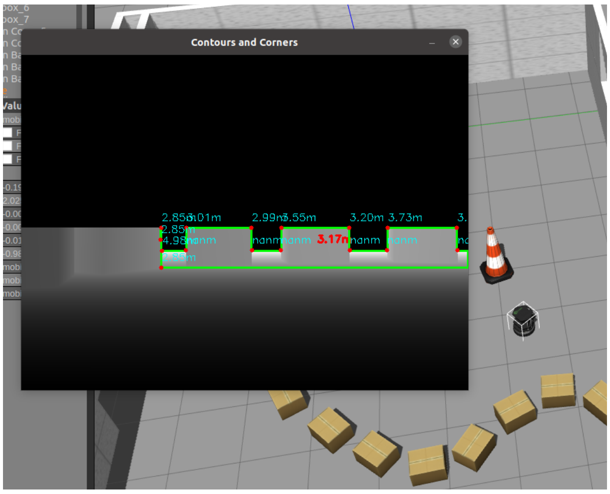

Fig.4. Boxes

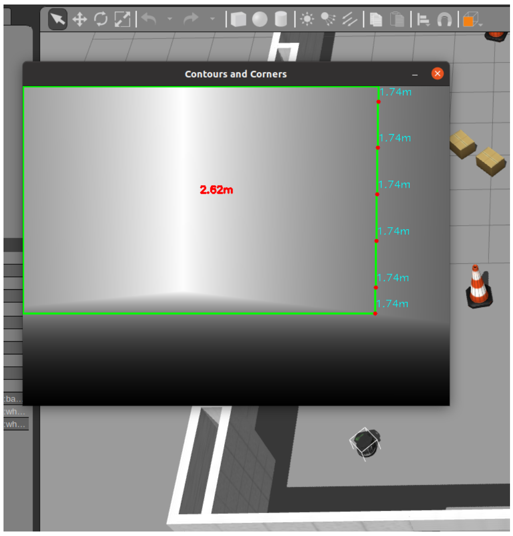

Fig.5. Wall edge

The images above are the visualisation of contour/corner detections and distance measurements. The green colour is used to indicate the contour and the red colour for corners. The distance to both contour centroid and corner points is displayed in metres.

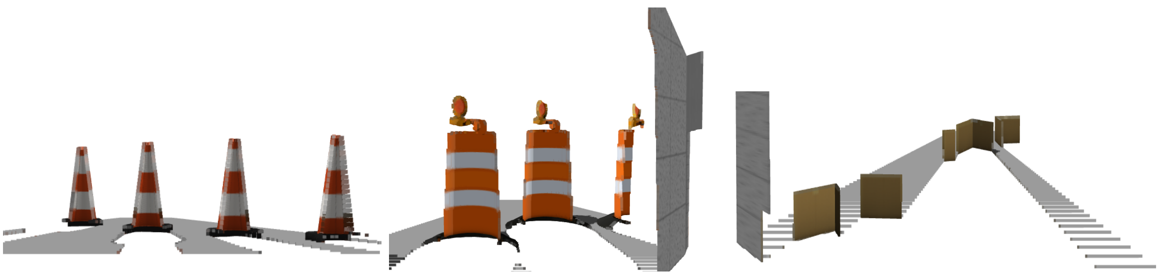

Fig. 6. 2D to point cloud

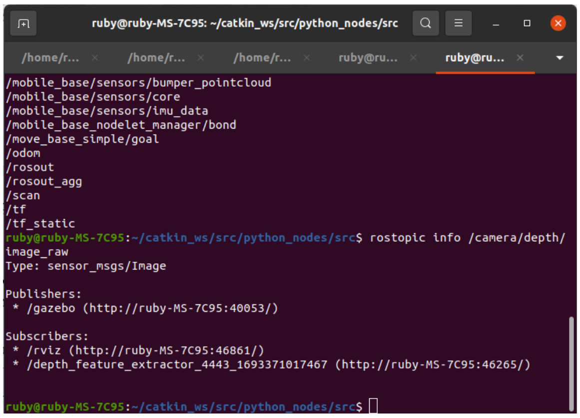

Fig.7. Publisher and subscriber

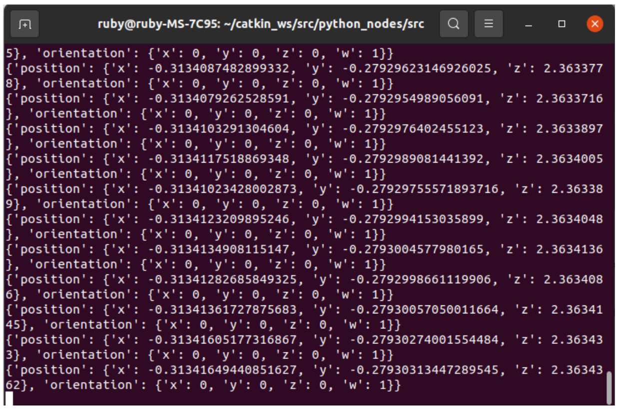

Fig.8. terminal

In the terminal, the 3D position of the pixel in the camera's coordinate frame is printed out with x, y, and z values. X and y are computed using the camera's focal lengths and optical centres. The z value is the actual depth value at the pixel. It represents the distance from the camera sensor to the object at that pixel.

Justification and evaluation

For the distance measurement evaluation, I tested the accuracy using the position of the turtlebot and the objects in the gazebo virtual environment.

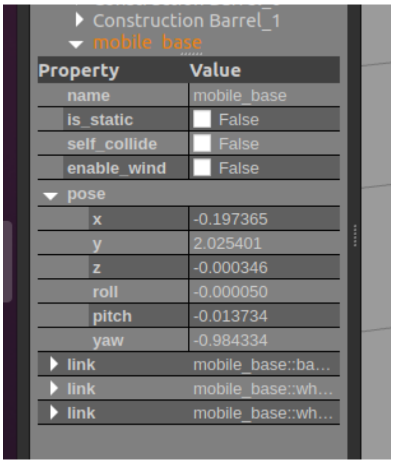

Fig.9. Position of turtlebot

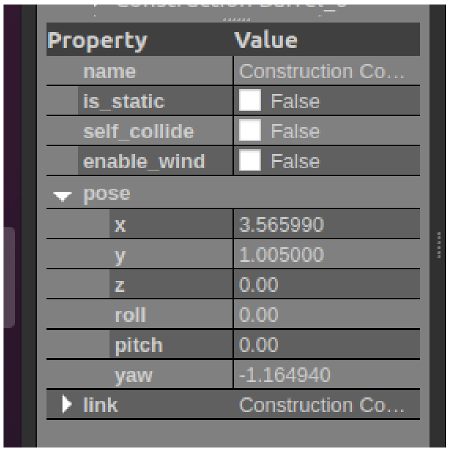

Fig.10. Position of a small cone

Cones

Dense: Small cones have tapering shapes and big cones have cylindrical shapes with a feature at the top. These shapes, especially small cones, distinctly vary from the base to the tip.

Easy to sense: The shape of the cones creates unique image in depth images. Their distinct shapes make them easy to sense and differentiate against boxes and flat walls.

Ideally unique: In this virtual environment, no object has the same shape as the cones. This makes them unique landmarks which can be easily detected and differentiated from other features.

Evaluation: I have done 10 trials comparing the distance displayed in the visualisation to the distance calculated within the Gazebo position. In 8 out of 10 trials, the differences were within 0.1 metres and cones were detected 100% in all 10 trials.

Boxes
Dense: Boxes have flat surfaces and right angles at the vertices. A depth image of the boxes has dense regions since the depth values change quickly at the box edges. This can be seen as a sharp transition, especially at the corners of the box.
Easy to sense: The uniform shape of boxes makes them easily differentiable in depth images due to sharp depth transitions at their edges and corners.
Ideally unique: In this virtual environment, no object has the same shape as the boxes which have flat surfaces with distinct right angles. This makes the boxes unique landmarks.
Evaluation: I have done 10 trials comparing the distance displayed in the visualisation to the distance calculated in Gazebo, just like cones. In 9 out of 10 trials, the differences were within 0.1 metres and boxes were detected 100%.

Wall corners

For the wall corners, while the implemented algorithm detected the wall edges successfully, it consistently failed to detect the wall corners as shown in Fig.6. I assumed that the wall corners would be a good landmark as the corner creates distinct angles in depth images. However, even changing the area threshold to detect and changing the binary threshold did not make the performance better. Therefore, I decided to use wall edges instead of wall corners as landmarks.

Wall edges
Dense: In indoor environments, such as the simulation world I'm using, wall edges provide a dense linear feature in depth images. These edges, especially in indoor environments, have a distinct change in depth values. The continuous nature of the edge offers a steady reference point, particularly emphasised by the sharp depth transitions.
Easy to sense: The structure of wall edges is easy to detect in depth images.
Ideally unique: In this virtual environment, wall edges have distinct lengths and structures which makes them unique landmarks. There are no other features that have continuous linear depth transitions in the environment.
Evaluation: In all 10 trials, the differences were within 0.1 metres and wall edges were detected 100% as well.

Trial / Difference in metres	Cones	Boxes	Walls
1	0.06	0.05	0.06
2	0.07	1.51	0.07
3	0.08	0.09	0.04
4	1.19	0.08	0.09
5	0.09	0.07	0.08
6	2.23	0.09	0.04
7	0.05	0.04	0.05
8	0.09	0.07	0.05
9	0.06	1.00	0.09
10	0.09	0.08	0.05

Table.1. Distance measurement trials

Conclusion

In conclusion, this project aimed to identify suitable features for robot localisation and mapping using depth-sensing cameras. Three potential landmarks were evaluated: cones, boxes, and initially, wall corners. However, wall corners fell short in contrast to my expectations and the wall edges were recognised as reliable landmarks. Therefore, in this virtual environment, cones, boxes, and wall edges could be used as landmarks for navigation and localisation. Their dense depth profiles, ease of sensing in depth images, and unique features make them suitable for landmarks.