Real-time Object Detection From LiDAR Point Clouds

Dong Lingzheng
2020533041
donglzh@shanghaitech.edu.cn

Zhang Jiajie
2023233216
zhangjj2023@shanghaitech.edu.cn

Abstract

Solar energy, as a crucial clean energy source, is playing an increasingly pivotal role in improving the human energy infrastructure. However, installing large solar panels in sun-rich desert areas still faces challenges like low manual efficiency and high costs. This project utilizes LiDAR to efficiently detect key information about objects during the installation process, such as the pose of the solar rack, the position and height of these supports, and the tilt angle of solar panels, enabling the automated installation of solar panels by robots. We conduct abundant experiments on various datasets of real solar racks provided by the company, demonstrating the correctness, efficiency, and generality of our method.

1. Introduction

In the pursuit of sustainable energy solutions, solar power has emerged as a pivotal contributor to a cleaner and more efficient global energy infrastructure. Expanding the utilization of solar energy, particularly in sun-rich desert regions, presents an opportunity for substantial energy generation. However, the installation of large-scale solar panels in such areas encounters challenges associated with low manual efficiency and high costs.

To address these challenges, this paper proposes a novel approach that leverages Light Detection and Ranging (LiDAR) technology for real-time object detection during the solar panel installation process. LiDAR, known for its precision and accuracy in capturing three-dimensional spatial data, is employed to efficiently gather key information about the environment. The primary focus lies in detecting the pose of the solar rack, determining the position and height of support structures, and assessing the tilt angle of solar panels.

The integration of LiDAR-based object detection enables the automation of solar panel installation, facilitated by robotic systems. This advancement not only contributes to increased operational efficiency but also mitigates the challenges associated with human labor in expansive and challenging terrains.

In this paper, we present the results of extensive experiments conducted on diverse datasets comprising real-world solar rack scenarios provided by the company. Our objective is to showcase the correctness, efficiency, and generality of the proposed method, emphasizing its potential to revolutionize the solar panel installation process. The outcomes of these experiments demonstrate the effectiveness of our approach in addressing the complexities inherent in real-world solar panel deployment, thereby laying the groundwork for a more streamlined and automated future in solar energy infrastructure.

2. State of the Art

2.1. Literature

"LiDAR-based detection, tracking, and property estimation: A contemporary review" [2], discusses advancements in Object detection, Person tracking, and Person property estimation (PPE), emphasizing the shift from image-based to sensor-based analysis, including infrared, depth cameras, and LiDAR sensors. It anticipates LiDAR-based 3D object detection becoming a prominent field, particularly in autonomous driving applications, offering a comprehensive survey of recent research in this area.

In "Instance Segmentation of LiDAR Point Clouds" [8], which presents a robust baseline method for instance segmentation in large-scale outdoor LiDAR point clouds, featuring a novel dense feature encoding technique for precise localization of small, distant objects, a simple yet effective approach for single-shot instance prediction, and strategies to address severe class imbalances.

In "GIS-based estimation of rooftop solar photovoltaic potential using LiDAR" [4], which utilizes LiDAR data merged with Geographic Information System techniques to identify optimal locations for solar panels on rooftops in the Georgetown area, considering criteria such as ground, aspect, slope, human factors, and high radiation levels. The findings provide insights into potential PV outputs, guiding energy policy, and future research on solar PV deployment.

In "Instant Object Detection in Lidar Point Clouds" [1], a novel method is introduced for object classification in continuously streamed Lidar point clouds obtained from urban environments. The proposed framework processes raw 3D point cloud sequences from a Velodyne HDL-64 Lidar, with a primary focus on identifying vehicles and pedestrians in the vicinity of the mobile sensor. The devised pipeline, tailored for outdoor 3D urban object recognition, initially segments the point cloud into the ground, short objects (low foreground), and tall objects (high foreground). Utilizing a unique two-layer grid structure, efficient connected component analysis is then performed on foreground regions to generate distinct groups of points representing various urban objects. Subsequently, depth images are generated from object candidates, and an appearance-based preliminary classification is conducted through a convolutional neural network. The classification is further refined with contextual features, taking into account expected scene topologies. The algorithm's performance is validated through testing on real Lidar measurements, comprising 1485 objects captured from diverse urban scenarios.

The review "Deep 3D Object Detection Networks Using LiDAR Data: A Review" [7] provides a comprehensive examination of challenges and methodologies in 3D object detection networks utilizing LiDAR data, encompassing an overview of the 3D detection task, LiDAR sensing techniques, deep 3D detection networks, challenges associated with LiDAR point cloud representations, evaluation metrics, algorithm performance on authoritative benchmarks, and insights into existing challenges and open issues.

In "LIDAR-based 3D object perception" [3], the authors present a LiDAR-based perception system for ground robot mobility, incorporating 3D object detection, classification, and tracking, demonstrated on the MuCAR-3 autonomous ground vehicle. The system efficiently navigates urban traffic and off-road scenarios through a unique combination of 2D and 3D data processing, achieving real-time operation at a 0.1s frame rate.

GndNet [5] proposes a novel end-to-end approach that estimates the ground plane elevation information and ground points segmentation in real time, achieving the state-of-the-art with best accuracy in time rate of 55Hz. The author augement the dataset derived from SemanticKITTI by CRF-based method to train the network. GndNet combines PointNet and Pillar Feature Encoding network to extract features of point clouds with occupancy grid generation, then regresses the ground height using a convolutional encoder-decoder network. It can estimate elevation of complicated grounds in real time and segment the point clouds into ground and non-ground categories, inspiring to obtain plane parameters in our work.

"Object Detection From a Few LIDAR Scanning Planes" [6] introduces a novel recognition method tailored for LIDARs, specifically focusing on a sparse number of detection planes. This approach proves advantageous, particularly in scenarios where the angular resolution is adequate, but the planes are widely spaced in the vertical direction. The methodology incorporates Fourier descriptors to characterize a scan plane and utilizes Convolutional Neural Networks for classification. Our method capitalizes on both time-varying shape information and contours derived from multiple scan planes when available. Demonstrating efficacy, the proposed method performs at least comparably to state-of-the-art algorithms in near-field situations while simultaneously extending the detection range. The evaluation involved tens of thousands of samples from extensive public datasets, encompassing separate assessments for far-field objects.

2.2. ROS Package

2.2.1 PCL

The Point Cloud Library (PCL) is a comprehensive open-source library within the Robot Operating System (ROS) environment, designed to facilitate the processing and analysis of 2D/3D point cloud data. PCL provides a rich set of functionalities for tasks such as point cloud filtering, segmentation, feature extraction, registration, and visualization. It supports a variety of point cloud data sources, including Lidar and depth cameras.

PCL is seamlessly integrated with ROS, allowing users to leverage its powerful tools for perception and manipulation tasks in robotic applications. It enables efficient handling and manipulation of point cloud data through a set of modular and extensible algorithms, making it a valuable resource for researchers, developers, and roboticists working on projects involving perception and environmental understanding.

The library supports both C++ and Python programming languages, providing flexibility for developers to choose the language that best suits their needs. With its extensive set of tools and algorithms, PCL in ROS contributes to advancing the capabilities of robotic systems by enabling them to interpret and interact with their environment through point cloud data.

2.2.2 Rviz

ROS Visualization (rviz) is a powerful 3D visualization tool that is an integral part of the Robot Operating System (ROS). It provides a user-friendly interface for visualizing various types of data generated within a ROS-based robotic system. With rviz, users can visualize and interact with data such as point clouds, robot models, sensor data, trajectories, and more.

Rviz allows users to customize their visualizations to better understand the robot's perception of its environment. It supports the display of information from different sensors and topics, enabling a comprehensive view of the robot's surroundings. Users can interactively manipulate the visualization, change perspectives, and gain insights into the robot's sensor data in real-time.

This visualization tool is particularly valuable for debugging, testing, and understanding the robot's behavior during development. It aids in the analysis of sensor data and the evaluation of algorithms, contributing to the overall understanding and improvement of robotic systems within the ROS framework.

3. System Description

3.1. Pipeline

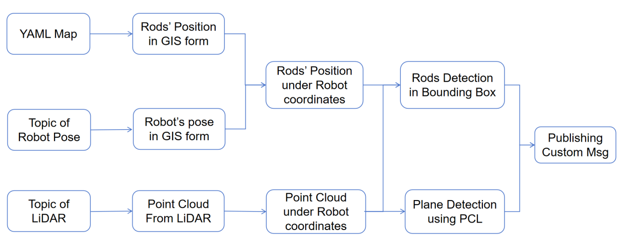


Figure 1. Overview of our framework

- 1. Read the longitude and latitude of origin, base link and rods from yaml map. Subscibe to pose of base link from driving group.
- 2. Transform the coordinates of rods under base link.
- 3. Subscribe to point cloud messages from 2 Lidars.
- 4. Transform the point clouds under base link according to the pose of the lidar w.r.t. the base link. Then perform point clouds fusion.
- 5. Start perception (Mainly use methods integrated in the PCL package)

  5.1 Using $statisticalremoval$ and $PassThrough$ filter methods to remove the outlier and noise and downsampling the points to accelerate the process.

  5.2 Using $Planemodel$ segmentation method to remove the ground and fit the solar panel, then get the parameters of the ground plane, the coordinate of the center of the panel, and the roll, pitch and yaw of the panel (maybe in long-lat form).

  5.3 Using $Linemodel$ segmentation method to fit the rods individually, given the prior knowledge of rods position and size, then get the coordinates of the top of the rods by averaging several top points.

  5.4 Compute the inclination degree of rods from the coefficients of line models and publish an alert if the inclination could cause the installation to fail.

- 6. Publish the custom messages of panel and rods infos (pose and errors) into a topic at a frequency of 10 Hertz.
- 7. Visualize the processed point cloud in Rviz lively, together with robot urdf, and use the marker to point out origin points, the detected solar panel and rods.

4. Experiments

4.1. Datasets

The radar dataset is collected by Bolight Company, named as "trans_2023-12-08-11-31-27.bag". To differentiate the frame id of each radar, we set the frame id of the left radar as "lidar1_link" and that of the right radar as "lidar2_link" and changes the file name by "3lidars.bag". The bag file contains all kinds of information such as point clouds, URDF infos, etc.. We only focus on the information of point clouds such as topics and messages, etc..

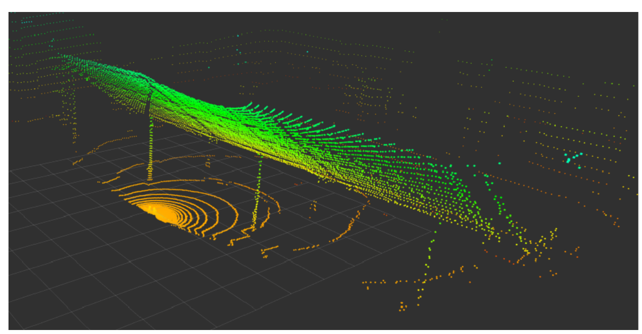


Figure 2. Original point cloud with solar panel and rods

4.2. Visualization Result

Fig3, Fig4, Fig5 and Fig6 are the pictures of the visualization. We use marker to mark each object detected in Rviz. The original point clouds are marked by white marker, the detected panel is marked by red marker, the prior rods are marked by yellow marker, and the detected rods are marked by purple marker.

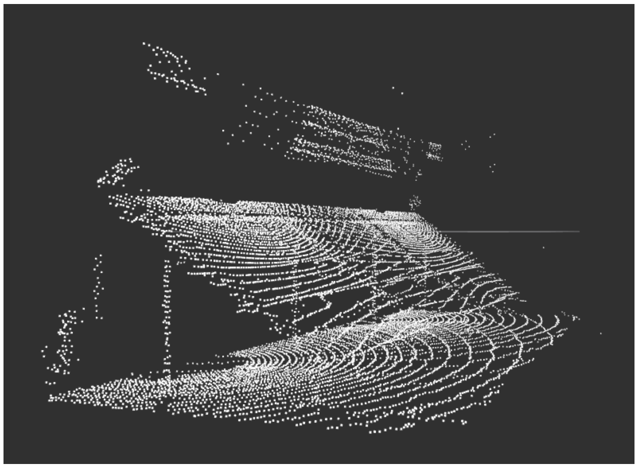


Figure 3. Point cloud of the whole scene

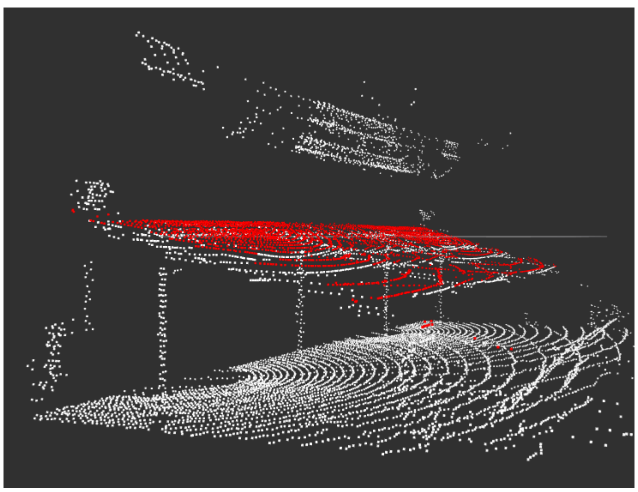


Figure 4. Point cloud of detected solar panel

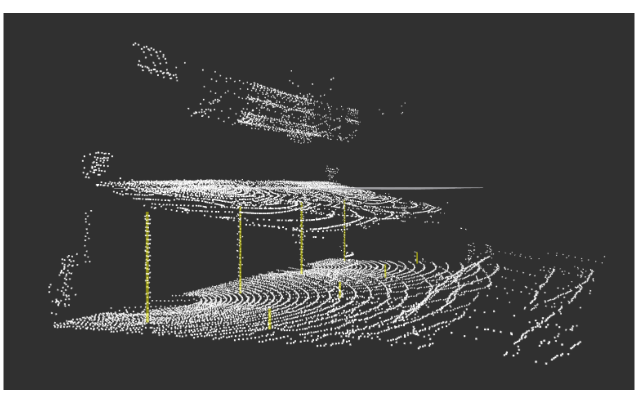


Figure 5. Point cloud of rods from prior knowledge

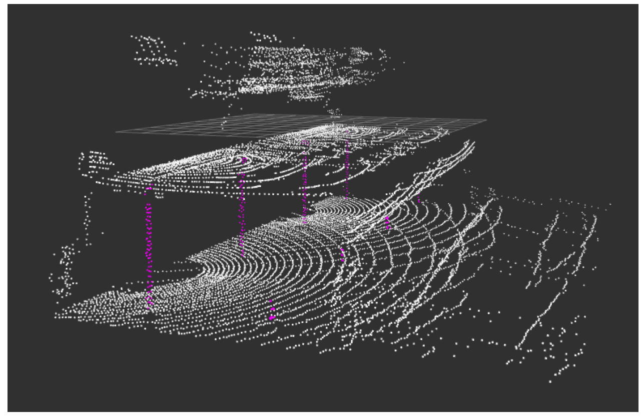


Figure 6. Point cloud of detected rods

4.3. Numerical Result

The custom messages of panel consists of the inclination of the panel (described by rall, pitch and yaw). The custom messages of rods consists of the position (described by x,y,z), inclination (described by rall and pitch) and errors (described by average error and max error) of eight rods. The custom messages of detected results are published in real time of 10Hz.

Belows are the format of custom messages.

```
src > detection > msg > Panel.msg
float32 roll
float32 pitch
float32 yaw
```

(a) Panel message

```
src > detection > msg > Rod.msg
float32 x
float32 y
float32 z
float32 roll
float32 pitch
float32 avg_error
float32 max_error
```

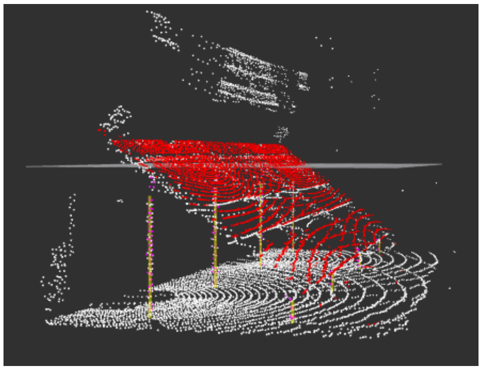
(b) Rod message

```
src > detection > msg > Rods.msg
Rod rod1
Rod rod2
Rod rod3
Rod rod4
```

(c) Rods message

In the custom messages, the first two lines describe the information of detected panel. The next 16 lines (2 lines for each rod) describe the information of detected rods, with the first 8 lines representing the rods from left side and the second 8 lines representing the rods from right side.

Belows are the numerical result of one frame in the bag file.



(a) Point cloud of one frame

```
The number of the points in the plane:5743
The inclination of panel: (36.7228,-2.02521).
The position of rod1: (-0.758714,-1.36179,-0.0963346).
The inclination of rod1: (-0.0157155,-0.00671406).
The position of rod2: (-5.14537,-1.70192,-0.0738262).
The inclination of rod2: (-0.0522938,0.0024156).
The position of rod3: (-0.557011,-3.70782,-2.04052).
The inclination of rod3: (-0.0157155,-0.00671406).
The position of rod4: (-4.9794,-4.02868,-2.00235).
The inclination of rod4: (0.180656,0.0024156).

The position of rod1: (-9.65672,-1.91556,-0.0876253).
The inclination of rod1: (-0.0172422,-0.00854469).
The position of rod2: (-14.0622,-2.15765,-0.144353).
The inclination of rod2: (-0.0033893,0.00195087).
The position of rod3: (-9.44206,-4.33501,-2.09989).
The inclination of rod3: (-0.0172422,-0.00854469).
The position of rod4: (-13.8631,-4.5502,-2.07316).
The inclination of rod4: (0.0577534,0.00195087).
```

(b) Custom messages of one frame

Figure 8. Custom messages of one frame

5. Conclusion

We have proposed an effective utilization of LiDAR to enhance the installation of large-scale solar panels. The real-time detection of critical information, including the precise positioning, height, and tilt angle of both solar panels and their supports, is achieved through a frequency of 10 Hertz message publication within the Robot Operating System (ROS). The integration of this technology enables robots to automate the installation of solar panels, ensuring accuracy in angles and positions. This standardized installation process not only improves the efficiency of solar panel deployment but also contributes to cost reduction by minimizing manual labor.

References

[1] Attila Börcs, Balázs Nagy, and Csaba Benedek. Instant object detection in lidar point clouds. *IEEE Geoscience and Remote Sensing Letters*, 14(7):992–996, 2017. 1

[2] Mahmudul Hasan, Junichi Hanawa, Riku Goto, Ryota Suzuki, Hisato Fukuda, Yoshinori Kuno, and Yoshinori Kobayashi. Lidar-based detection, tracking, and property estimation: A contemporary review. *Neurocomputing*, 506:393–405, 2022. 1

[3] M Himmelsbach, A Müller, Thorsten Luettel, and Hans-Joachim Wuensche. Lidar-based 3d object perception. 10 2008. 2

[4] Zulkiflee Abd Latif, Nurul Ain Mohd Zaki, and Siti Aekbal Salleh. Gis-based estimation of rooftop solar photovoltaic potential using lidar. In *2012 IEEE 8th International Colloquium on Signal Processing and its Applications*, pages 388–392, 2012. 1

[5] Anshul Paigwar, Özgür Erkent, David Sierra-Gonzalez, and Christian Laugier. Gndnet: Fast ground plane estimation and point cloud segmentation for autonomous vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2150–2156, 2020. 2

[6] Zoltan Rozsa and Tamas Sziranyi. Object detection from a few lidar scanning planes. *IEEE Transactions on Intelligent Vehicles*, 4(4):548–560, 2019. 2

[7] Yutian Wu, Yueyu Wang, Shuwei Zhang, and Harutoshi Ogai. Deep 3d object detection networks using lidar data: A review. *IEEE Sensors Journal*, 21(2):1152–1171, 2021. 2

[8] Feihu Zhang, Chenye Guan, Jin Fang, Song Bai, Ruigang Yang, Philip H.S. Torr, and Victor Prisacariu. Instance segmentation of lidar point clouds. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9448–9455, 2020. 1