

## 安価な深度カメラを用いた駐車車両位置推定手法の検討

WEN Jun† 田口 亮†

名古屋工業大学†

## 1. はじめに

近年、3次元点群からの車両検出技術の発展に伴い、自動運転やスマートトラフィックなどへの応用が注目されている。しかし、従来の手法は高価なLiDARで得られた点群データに依存しており、大規模な応用には制約がある。例えば中古車やレンタカーなどの車両を自動外観検査するロボット<sup>[1]</sup>を想定すると、より安価な深度カメラを用いて駐車車両の3次元位置、姿勢が推定できることが望ましい。そこで本研究では、深度カメラを用いて駐車場の車両位置推定を実現する手法を提案する。

## 2. システムの概要

## 2.1 車両点群の抽出

RGB-DカメラからRGB画像とデプス画像を取得する。YOLOv5<sup>[2]</sup>を用い、RGB画像から車両を検出する。得られたバウンディングボックスの左上隅と右下隅の座標をSAM(Segment Anything Model)<sup>[3]</sup>の入力とし、車両領域をピクセル単位で抽出する。得られた領域に対応する車両以外のピクセルを除去する。RGB画像とデプス画像の対応を用いて、車両ピクセルに対応する3次元点群を抽出する。

## 2.2 車両の3次元位置・姿勢の推定

得られた車両点群は車両の一部である。この点群を学習済みのVoxelNet<sup>[4]</sup>に入力することで、車両全体のバウンディングボックスを推定する。ただし、VoxelNetの学習には走行時のLiDARデータが用いられており、駐車車両をRGB-Dカメラで撮影した際に得られる点群データとは傾向が異なる。具体的には、後者は前者よりも密な点群が得られるが、鏡面反射によるノイズが含まれる場合もある。そのためRGB-Dカメラで得られた点群をVoxelNetに入力しても正しい姿勢が得られない。そこで、距離点群を高さに応じて複数の部分に分割し、それぞれをVoxelNetの入力とし、車両のバウンディングボックスを推定する。得られた複数のバウンディングボックスをクラスタリングにより1つにまとめる。具体的



図1 仮想環境

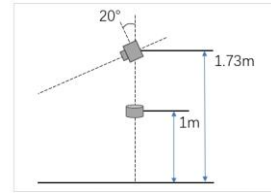


図2 配置図

場所：名古屋工業大学15館  
前駐車場  
天気：曇り  
カメラ：Intel RealSense  
Depth Camera D455



図3 屋外駐車場とカメラの概要

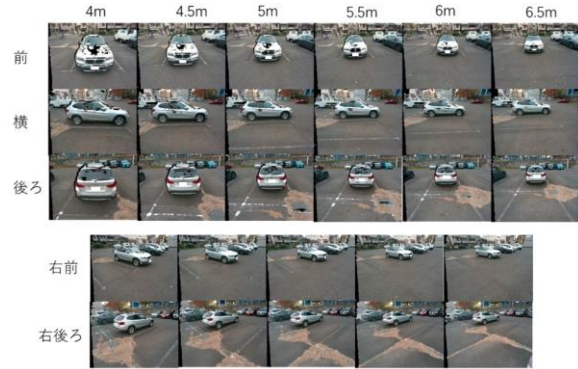


図4 屋外駐車場

にはボックスの中心点をk-means法によりクラスタリングし、シルエット係数で最適なクラス数を推定する。そして平均scoreが最も高いカテゴリーの平均をバウンディングボックスの重心とする。また、サイズ(幅、高さ、奥行き)と向き(yaw)もそれぞれ先ほど得られたクラス内で平均する。

## 3. 実験環境

## 3.1 仮想環境

図1に実験に用いる仮想環境、図2にRGB-Dカメラの姿勢を示す。仮想環境にはLiDARとRGB-Dカメラを搭載したロボットと一台の車を配置する。図2の上がカメラ、下がLiDARである。目標車両を原点とし、距離4m~6.5mの範囲を0.5m刻み、周囲5°刻みにカメラを移動し、合計222の位置から撮影する。

## 3.2 屋外駐車場

RGB-Dカメラは赤外のパターン光を使用して深度画像を取得するため、直射日光の下では撮影

A Method for Estimating Location of Parked Vehicles using Inexpensive Depth Camera

† Jun Wen and Ryo Taguchi, Nagoya Institute of Technology

が難しい。そのため、建物の陰で直射日光の影響を受けないような場所と時間を図 3 に示すように設定した。固定スタンドを使用して、カメラを仮想環境と同様に配置した。図 4 に示すように、目標車両の前方から後方までを 45° 刻み、距離を 4m から 6.5m まで 0.5m 刻みでカメラを移動し合計 28 箇所から撮影した。図 4 に示す画像は RGB 画像を深度カメラの座標系に変換した結果である。車両の反射やオクルージョンにより距離値が計測できなかった箇所は、黒く表示されている。

#### 4. 実験結果

車両の正しいバウンディングボックスと、推定したバウンディングボックスを用いて、俯瞰視点で IoU を計算する。また、各撮影位置 IoU から平均 IoU を計算する。仮想環境では、LiDAR で得られた点群を用いた場合の結果と提案手法を比較する。

図 5 に仮想環境で LiDAR を用いた場合の結果を示す。灰色の点はモデルが結果を出力しなかったことを示し、この場合、IoU は 0 である。赤い点は IoU が 0.4 未満であり、検出失敗を示す。オレンジの点は IoU が 0.4 以上 0.6 未満であり、車両の検出が困難であることを示す。緑の点は IoU が 0.6 以上であり検出成功を示す。図中の赤字で書かれた値が平均 IoU である。図 6 に仮想環境でデプス画像を用いた提案手法の性能を示す。出力の平均 IoU が LiDAR を用いた結果よりも高いことから、提案手法は仮想環境で優れた性能を発揮していることが示されている。

図 7 に実環境での実験の結果を示す。青い矩形が車両の位置を示しており、重心がグラフの原点(0, 0)、車両の前方はグラフの下を向いている。平均 IoU はやや仮想環境の結果よりも低いですが、実車両では反射があることを考慮すると、提案システムは良好な性能を発揮していると言える。

#### 5. まとめ

本研究では、深度カメラを用いて車両位置を推定できる手法を提案した。提案手法では、まず深層学習に基づく車両検出手法に基づき車両だけの点群を抽出する。その後、高さ方向にダウンサンプリングして得られた複数の点群を学習済みの VoxelNet に入力して複数の推定結果を取得する。最後に、複数の予測位置をクラスタリングにより 1 つにまとめる。仮想環境と実環境での実験により提案手法の有効性を示した。本

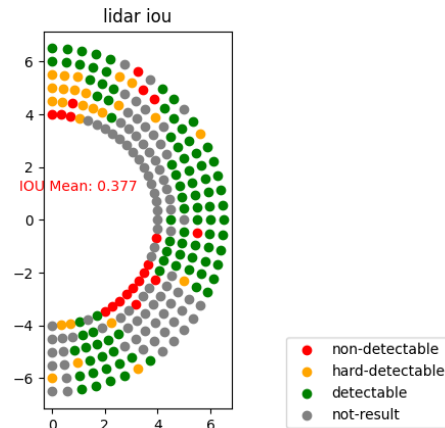


図 5 仮想環境 (Lidar) の結果

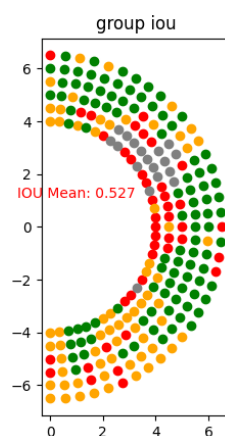


図 6 仮想環境 (デプス画像) の結果

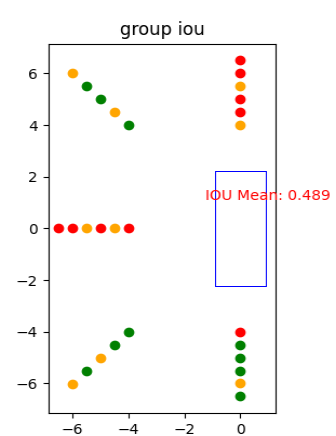


図 7 実環境での結果

研究では車以外の障害物が含まれないような環境で実験を行った。車止めやカラーコーンなど、他の障害物がある環境での検証は今後の課題である。

#### 参考文献

- [1] 史智文, 田口亮, 「深層学習による物体認識と VSLAM を併用した屋外環境での駐車車両位置推定」, 第 37 回 日本ロボット学会 学術講演会, 2K2-05, 2020.
- [2] Ultralytics, "YOLOv5", <https://ultralytics.com/yolov5>
- [3] Kirillov, Alexander, et al. "Segment anything," arXiv:2304.02643 (2023).
- [4] Zhou, Yin, and Oncel Tuzel. "Voxelnet: End-to-end learning for point cloud based 3d object detection," Proc. of CVPR 2018.