

# 単眼カメラから得られる深度情報を用いた 深層強化学習に基づく移動ロボットの自律走行システム

○横山 光希 (明治大学), 森岡 一幸 (明治大学)

## Mobile Robot Navigation System Based on Deep Reinforcement Learning Using Depth Estimation from a Monocular Camera

○ Koki YOKOYAMA (Meiji University), and Kazuyuki MORIOKA (Meiji Univsity)

**Abstract :** The purpose of this study is development of an autonomous mobile robot navigation system based on deep reinforcement learning with a monocular camera. Monocular camera-based navigation can simplify robot system configuration without LiDARs. The system requires the input data as states of reinforcement learning that include the range data around the robot. In this paper, the input range data is estimated from a monocular camera instead of 2D-LiDAR. Then, we conduct two experiments and evaluate autonomous navigation by the proposed system in the real world. The results show the autonomous navigation was achieved according to camera image-based states.

### 1. 緒言

自律走行はロボットにおける数ある課題の中でも、基礎的かつ重要なものである。主に自律走行の課題として自己位置推定と経路計画が挙げられる。広く知られる手法として 2D-LiDAR とグリッドマップで行うものがある。近年では深層強化学習により得られた行動方策を用いて実環境での自律走行を実現する事例も報告されている [1][2]。深層強化学習は現在の状態から次のアクションを適切に選択する方策を、多層のニューラルネットワークにより表現する方法である。また、KITTI や Cityscapes などの画像データセットの普及により、画像ベースの走行システムの研究開発が盛んに行われている。特に単眼カメラから深度情報を取得する研究で成果を挙げている [3]。

本研究では、深層強化学習による行動獲得と画像ベースの状態表現に基づいた、移動ロボットの自律走行システムを提案する。特に LiDAR を一切使用せず、単眼カメラに基づく簡易なシステム構成による移動ロボットの開発を目指している。それにより、例えば現在ほとんどの人が保持しているスマートフォンを、ロボットのセンサおよびコンピュータとして取り付けるだけで自律走行が可能となる (Fig.1)。つまり LiDAR よりも安価で手に入りやすいセンサによりロボットシステムが構成できるようになる。さらには、深層強化学習ベースの自律走行にすることにより、事前に詳細な環境地図を用意せずに自律走行可能なシステムを目指している。本研究は、ハードウェアおよびソフトウェアの双方において、ロボットシステムを単純化、高度化する試みである。

本稿ではその初期段階として、単眼カメラから得られる

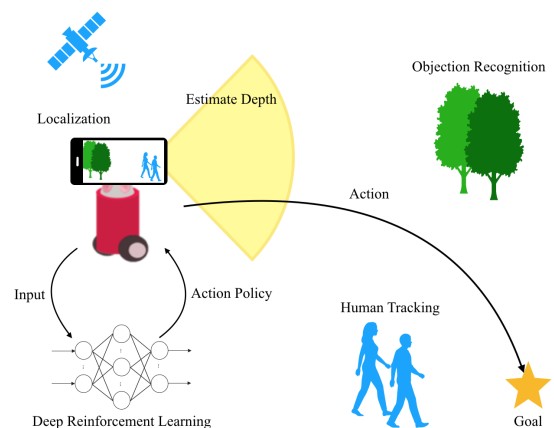


Fig. 1: Concept of the proposed system

深度情報を強化学習モデルの入力の一つとして用いた、ロボットの屋内自律走行の実現を目的とする。そして自律走行実験を行い、システムの有用性を確認する。

### 2. 提案システム

#### 2.1 提案システム概要

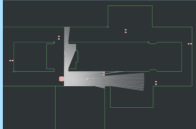
本章では開発したシステムの概要および構成要素について述べる。本稿の提案システムは主に以下の3つの要素で構成されている。

- (1) 深層強化学習に基づく移動ロボットの行動獲得
- (2) 単眼カメラによる深度情報の取得
- (3) 2D-LiDAR による自己位置推定

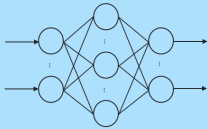
上記3つの要素を構成することにより、単眼カメラから取得できる深度情報を入力として用いた深層強化学習に基づ

# Deep Reinforcement Learning for Robot Action

learning on the simulator




Double Deep Q-Network



Applying the learning model


## Navigation System

Monocular camera for get distance

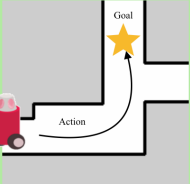


Distance of around

2D-LIDAR for localization



Distance and Angle to Goal



Goal

Action

## 2.2 システム構成

深層強化学習の環境は独自で開発した 2D-LiDAR ベースのシミュレータを利用する。このシミュレータは二次元平面でのロボットの行動方策を学習する用途に特化しており、一般的なロボット用シミュレータと比較し高速な学習が可能である [1]。このシミュレータでは、実環境を模した環境を構築し、また静的物体だけでなく人間を模した動的物体も作り出すことができる。学習自体は 2D-LiDAR による距離測定を用いるが、実際の自律走行はカメラから得られる深度情報を用いる。我々はこれまでに 2D-LiDAR により得られた距離情報を状態として学習した行動方策を用いて自律走行が可能なることを報告してきた [2]。2D-LiDAR ベースの学習で適切な行動選択が可能になるモデルが獲得できることが明らかであり、実際の走行時にのみカメラの深度情報を学習済みモデルに入力することで、カメラペー

[illegible]

### 3. 深層強化学習によるロボットの行動獲得

### 3.1 学習モデル

### 3.2 狀態

(a) 10 scan ranges

(b) distance and angle to goal

- 182 -

$$s_t = (sl_t^{norm}, sd_t^{norm}, sa_t^{norm}) \quad (1)$$

### 3.3 行動

強化学習モデルの出力として得られるロボットの行動  $a_t$  は3つあり、それぞれ並進速度  $v_l$  と角速度  $v_a$  によって Eq.2 のように構成される。3つの行動は直進・左折・右折を意味する。

$$\begin{aligned} a_t &= (v_l, v_a) \\ &= ((0.4, 0.0), (0.25, 0.5), (0.25, -0.5)) \end{aligned} \quad (2)$$

### 3.4 報酬

学習の際に用いる報酬  $r_t$  は Eq.3 のとおりである。目的地に到着した場合には正の報酬を、壁や人といった障害物への衝突とロボット走行時に前方 50cm に障害物がある場合には負の報酬を与える。また、時刻  $t$  の時点の目的地までの距離  $sd_t$  とその前の時刻  $t-1$  の時点での目的地までの距離  $sd_{t-1}$  の差も報酬とする。目的地に接近する場合に正の報酬が与えられることになる。

$$r_t = \begin{cases} 10.0 & \text{if arrive}(sd_t < 0.5) \\ -10.0 & \text{elif collision} \\ -0.5 & \text{elif } sl_t < 0.5 \\ 4(sd_{t-1} - sd_t) & \text{else} \end{cases} \quad (3)$$

## 4. 単眼カメラによる深度推定と学習モデルへの適用

### 4.1 単眼カメラによる深度情報の取得

我々は単眼カメラから深度情報を取得する手法として struct2depth[3] を用いる。github で公開されているコードと web 上で配布されている学習済みモデルを基に、単眼カメラの映像からリアルタイム深度情報を取得できるように実装した。実際に我々の環境で深度情報を取得したものが Fig.5 である。

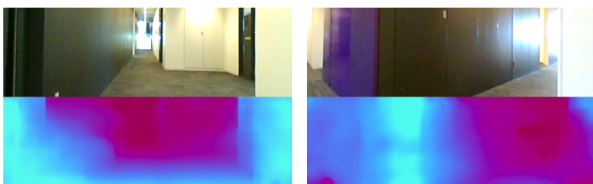


Fig. 5: Results of the depth estimated with a camera

### 4.2 深度情報から学習モデルの入力値への変換

深度情報を学習モデルの入力値として利用するために以下の処理を行う。その処理の流れを Fig.6 に示す。

- (1) 深度情報からカメラ座標系への変換
- (2) カメラの高さの深度情報を取得
- (3) 極座標系への変換
- (4) 入力値の作成
- (5) 実距離へのスケーリング

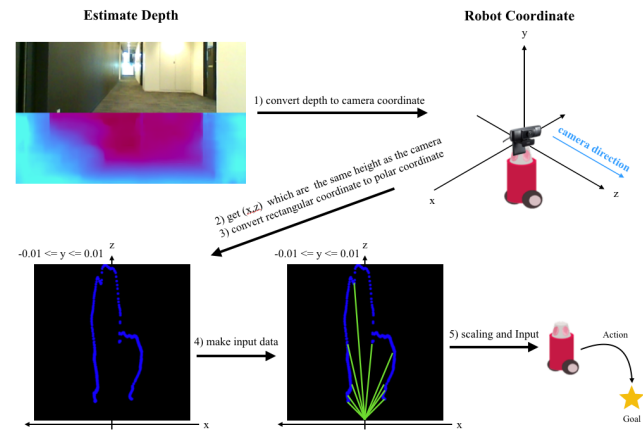


Fig. 6: the processing flow from depth to input state

2D-LiDAR ベースの入力値と同様に、カメラ画像の深度情報のうち床面に平行な平面での二次元形状を取得する必要がある。今回のシステムでは、カメラの高さの深度情報から二次元形状を取得する。そこでまずは深度情報を Fig.6 に示す xyz 空間であるカメラ座標系へと変換する。

カメラ座標系に変換したのちカメラの高さの平面での二次元形状を取得する。この座標系において原点はカメラであるので、カメラの高さの二次元形状を取得するためには  $y = 0$  の  $x, z$  座標を取得してくれば良い。また今回は少し幅を持たせるために  $-0.01 \leq y \leq 0.01$  の範囲の  $x, z$  座標を利用する。

実際の 2D-LiDAR の情報は角度と距離によって構成されている。取得してきた  $x, z$  座標系から極座標系へと変換することでそれに合わせることができる。また  $-0.01 \leq y \leq 0.01$  の範囲の極座標系の角度を計算したところ、 $45^\circ \leq \theta \leq 135^\circ$  の範囲に収まっており、本稿で利用しているカメラは前方  $90^\circ$  の二次元形状を取得できる。これは 3. 章で述べた深層強化学習の状態にて、ロボット前方  $90^\circ$  の形状を使用したものと合致している。その後得られた  $90^\circ$  の二次元形状を学習時と同様に 10 個のセグメント

に分けて、各セグメントの最小値を入力の情報とする。

最後に実距離へのスケーリングを行う。実環境の特徴的なランドマークを用いてスケーリングのパラメータを求める。ここでは、両端が壁になっている環境にて画像から得られた二次元形状を実際の距離情報と比較することでスケーリングの基準を算出した。

## 5. 実験

### 5.1 実験概要

この実験ではカメラから取得した深度情報を用いる自律走行と、2D-LiDAR により取得する距離情報を用いる自律走行の2つで行い比較をする。またこの実験では人や物といった障害物は存在していない。実験として straight path, bending path の2つの状況を想定して行い、走行および走行の成功確率の確認を行う。今回はそれぞれ別の状況を10回ずつ行い、成功確率を導き出す。

### 5.2 実験結果・考察

実際にロボットが走行した経路の一例を Fig.7 にて示す。どちらのシステムにおいても目的地に到達していることがわかる。しかし 2D-LiDAR を用いたシステムの方が滑らかな経路で走行している。これはカメラから得られる二次元形状は、学習時に用いた 2D-LiDAR で得られる距離情報と比較して精度は低く、適した行動が得られない場合が多いためだと考えられる。また現在のシステムにおいてロボットの行動選択の周期を取得したところ 2D-LiDAR の時は約 0.04 秒に一度、単眼カメラの時は約 0.10 秒に一度であり、この結果も走行安定性の原因だと考えられる。

また、成功確率として 2D-LiDAR を用いた時の straight path は 100%、bending path は 80% であり、カメラを用いた時は straight path、bending path どちらも 70% であった。bending path では右折する際に、曲がり角にぶつかることで失敗することを確認している。カメラ利用時の straight path ではまっすぐに進まず途中で右折してしまうことがある。これはカメラから得られる深度情報が安定性に欠けるため、ロボットが分かれ道に差し掛かった際に、右の方が開いていると判断し右折をして、そのまま復帰できなくなっていると考えられる。文献 [2] などではロボット前方 180° の形状を用いた行動学習をして 2D-LiDAR を用いたため走行経路が安定していたが、今回のように前方 90° にした場合は特に右折の際は 2D-LiDAR を用いた場合も成功率が低下している。前述の通りカメラから得ら

れた二次元形状の精度の低さに起因する性能低下も考えられるが、走行失敗の大きな原因はカメラを用いたことよりも学習および走行時の観測範囲が狭いことにあると思われる。

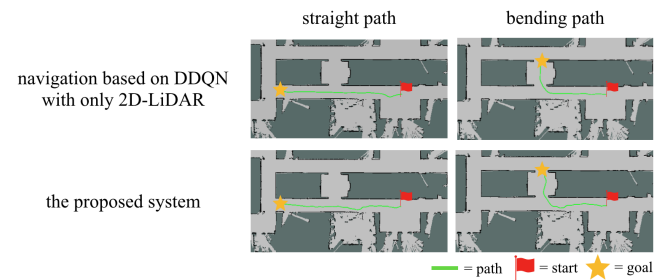


Fig. 7: navigation experiment

## 6. 結言

本稿では LiDAR を用いない自律走行システムについての概要を述べ、その初段階として単眼カメラから得られる深度情報を用いた強化学習ベースの自律走行システムを開発した。そして自律走行実験を行いその精度を確認した。

今後の課題としては以下のことが挙げられる。

- ・ 障害物ありの走行実験, 長距離走行実験
- ・ 深度推定精度向上のための on-line refinement
- ・ 2D-LiDAR を用いない自己位置推定の検討
- ・ 時系列データを扱うための学習アルゴリズムの変更

## 参考文献

- [1] Yuki Kato, Koji Kamiyama, and Kazuyuki Morioka, "Autonomous robot navigation system with learning based on deep q-network and topological maps", in Proc. of 2017 IEEE/SICE International Symposium on System Integration (SII), pp.1040 – 1046, 2017.
- [2] Yuki Kato, Kazuyuki Morioka, "Autonomous Robot Navigation System Without Grid Maps Based on Double Deep Q-Network and RTK-GNSS Localization in Outdoor Environments", in Proc. of 2019 IEEE/SICE International Symposium on System Integration (SII), pp.346-351, 2019.
- [3] V. Casser, S. Pirk, and R. Mahjourian, and A. Angelova, "Depth Prediction without the Sensors: Leveraging Structure for Unsupervised Learning from Monocular Videos", Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19), 2019
- [4] Hado Van Hasselt, Arthur Guez, and David Silver, "Deep Reinforcement Learning with Double Q-Learning", Association for the Advancement of Artificial Intelligence (AAAI), pp.2094-2100, 2016