CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Electrical Engineering

# MASTER'S THESIS

Zdeněk Rozsypálek

## Brick Detection for MBZIRC Competition

**Department of Control Engineering**

Thesis supervisor: **PhD. Petr Štěpán**

## Prohlášení autora práce

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze dne ............................ ..............................................

podpis autora práce

## Author statement for undergraduate thesis

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses.

Prague, date ............................ ..............................................

signature

# ZADÁNÍ DIPLOMOVÉ PRÁCE

## I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Rozsypálek** Jméno: **Zdeněk** Osobní číslo: **457216**

Fakulta/ústav: **Fakulta elektrotechnická**

Zadávající katedra/ústav: **Katedra řídicí techniky**

Studijní program: **Kybernetika a robotika**

Studijní obor: **Kybernetika a robotika**

## II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

**Detekce cihel pro soutěž MBZIRC**

Název diplomové práce anglicky:

**Brick detection for MBZIRC competition**

Pokyny pro vypracování:

Seznam doporučené literatury:

[1] Himmelsbach, Michael, et al. "LIDAR-based 3D object perception." Proceedings of 1st international workshop on cognition for technical systems. Vol. 1. 2008.
[2] Dou M., Guan L., Frahm JM., Fuchs H. (2013) Exploring High-Level Plane Primitives for Indoor 3D Reconstruction with a Hand-held RGB-D Camera. In: Park JI., Kim J. (eds) Computer Vision - ACCV 2012 Workshops. ACCV 2012. Lecture Notes in Computer Science, vol 7729. Springer, Berlin, Heidelberg
[3] Ma, L., Kerl, C., Stückler, J., & Cremers, D. (2016, May). CPA-SLAM: Consistent plane-model alignment for direct RGB-D SLAM. In 2016 IEEE International Conference on Robotics and Automation (ICRA) (pp. 1285-1291). IEEE.

Jméno a pracoviště vedoucí(ho) diplomové práce:

**RNDr. Petr Štěpán, Ph.D., Multirobotické systémy FEL**

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **14.01.2020** Termín odevzdání diplomové práce: ______________

Platnost zadání diplomové práce:
**do konce letního semestru 2020/2021**

| RNDr. Petr Štěpán, Ph.D.<br>podpis vedoucí(ho) práce | prof. Ing. Michael Šebek, DrSc.<br>podpis vedoucí(ho) ústavu/katedry | prof. Mgr. Petr Páta, Ph.D.<br>podpis děkana(ky) |
|---|---|---|

## III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací.
Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

.

Datum převzetí zadání Podpis studenta

## Acknowledgements

I would like to express my appreciation to Ing. Tomáš Petříček for his valuable and constructive suggestions during the planning and development of this thesis. I would also like to thank to the Department of Cybernetics of the Czech Technical University and to Michal Němec for the provided hardware. Finally, I wish to thank my family for support throughout my study.

## *Abstract*

Bakalářská práce se zaměřuje na ovládání *solid-state* lidarů s omezeným počtem natáčecích paprsků. Kromě plánování směrů paprsků se práce věnuje i rekonstruování 3D mapy z řídkých měření těchto lidarů. V práci se pro rekonstruování a plánování používají hluboké neuronové sítě. Plánovací část využívá *reinforcement learning* metody pro trénink neuronových sítí. Bylo vytvořeno trénovací prostředí implementující framework pro trénování *reinforcement learning* agentů. Za pomocí stochastických metod se podařilo navrhnout agenta, který nabízí dostatečnou škálovatelnost a překonává náhodný plánovač.



## *Abstrakt*

This Bachelor's thesis aims at control of the solid-state lidar sensor with a limited number of steerable rays. Besides planning of directions of the rays, the thesis is also devoted to creating dense 3D maps from sparse measurements. The thesis uses deep neural networks for planning the rays and reconstructing the dense maps. Planning part exploits the reinforcement learning concept for training of the neural network. An environment implementing a framework for training of reinforcement learning agents was created. The agent proposed in this thesis is using stochastic methods to achieve a sufficient scalability in the challenging environment.



**Keywords:** Lidar, reinforcement learning, deep neural network, 3D mapping, voxel map.

# Contents

# List of Figures

# 1 Introduction

Lidar sensors offer an accurate distance measurement, which can be used for mapping surrounding space. There is much utilization of volumetric space reconstructions in different fields. For example, the lidar sensors are nowadays essential equipment for a large variety of autonomous vehicles. The sensor can help autonomous vehicles to orient itself in an environment. One of the most significant issues which prevent a broader implementation of these sensors is a relatively high price. Breakthrough in this field is a solid-state lidar. These lidars do not have moving parts, and their price should be circa hundreds of dollars [1]. Solid-state lidar can send a limited number of rays in chosen directions per timestamp. Zimmermann et al. [2] proposed a mapping agent which creates dense reconstructions from sparse measurements. They also proposed prioritized greedy planning for choosing the directions of these rays.

The objective of this thesis is to apply reinforcement learning (RL) methods to learn planning of the rays and contribute to the methods of controlling these sensors. RL is a field of study based on concepts of behavioral psychology, especially the trial and error method, and has in recent years experienced a rapid development due to the growth of computational power and neural networks improvement. Richard Sutton has made a helpful summary of RL concepts in his book [3]. One of the biggest achievements was playing Atari games by a RL agent without any prior knowledge of the environment [4]. Soon after was introduced a RL agent, able to solve simple continuous problems such as balancing inverse pendulum on a cart. Today state-of-the-art methods can solve complex problems with infinite action spaces. Although these methods reach the great success, they still suffer from a lack of sample efficiency - they need for training a lot of interactions with the environment. This inefficiency makes creating an agent controlling lidar very challenging, since training large neural networks is very time-consuming.

The agent is divided into two parts - mapping and planning. The mapping part should create the best possible reconstruction from sparse measurements, while the planning part is focused on picking rays that will maximize reconstruction accuracy. Agents are trained using a publicly available dataset which contains drives of a car equipped with Velodyne lidar [5]. Theoretical background of the RL is discussed in the first part of this thesis. In the second part are methods from the first part used to solve the Lidar-gym environment [6].

# 2 MBZIRC

BLAH BLAH BLAH

# 3 Algorithms

BLAH BLAH BLAH

# 4 Experiment

BLAH BLAH BLAH

# 5 Conclusion

First, we overviewed reinforcement learning concepts and described several methods which help convergence of the learning process. Then, we addressed the challenging multi-dimensional control task of selecting depth-measuring rays for the 3D mapping. Various agents and model architectures were implemented and compared. All deterministic agents performed poorly in this specific task. The stochastic agent successfully outperformed the random planner. Action space size and time-complexity were two major blockers during the training. None of the trained RL agents can compete with the prioritized greedy planner proposed by Zimmermann et al. [2].

## 5.1 Future work

We propose further experiments with an agent, which stands between the simple and the extended stochastic agent. The extended stochastic agent has the action space consisting of 60 real numbers (15 rays with azimuth and elevation and for each alpha, beta parameters). That is very likely too much for the network architecture used in the experiments. On the other hand, when only one distribution is outputted for all rays, it does not allow the agent to create an advanced policies, because the Beta distribution considered in this thesis is always unimodal. A solution could be to output for example three different distributions, each describing five rays. That would allow agent to output a density function with multiple local maxima.

Another improvement could be achieved by adjusting the neural network architecture. Especially splitting the network into two subnetworks before the output or different types of merging the input layers can have a significant impact on performance. Finally, the reinforcement learning agent can be almost always improved by a better reward function, but we find very difficult to improve the existing reward function.

# References

[1] Evan Ackerman. Quanergy announces $250 solid-state lidar for cars, robots, and more. `https://spectrum.ieee.org/cars-that-think/transportation/sensors/quanergy-solid-state-lidar`, 2016.

[2] K. Zimmermann, T. Petricek, V. Salansky, and T. Svoboda. Learning for Active 3D Mapping. *ArXiv e-prints*, August 2017.

[3] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*, volume 2. The MIT Press, 2012.

[4] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, and Georg et al. Ostrovski. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[5] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*, 2013.

[6] Zdeněk Rozsypálek. Lidar-gym, training environment in openai interface. `https://gitlab.fel.cvut.cz/rozsyzde/lidar-gym`, 2018.

# Appendix A CD Content

In Table 1 are listed names of all root directories on CD.

| **Directory name** | **Description** |
|---|---|
| thesis | the thesis in pdf format |
| ctu_thesis | latex source codes |
| lidar-gym | OpenAI gym environment |

Table 1: CD Content

# Appendix B List of abbreviations

In Table 2 are listed abbreviations used in this thesis.

| **Abbreviation** | **Meaning** |
|---|---|
| **CNN** | Convolutional neural network |
| **DDPG** | Deep deterministic policy gradients |
| **DQN** | Deep Q-learning |
| **MDP** | Markov decision process |
| **RL** | Reinforcement learning |
| **ReLu** | Rectified linear unit |
| **TD** | Temporal difference |

Table 2: Lists of abbreviations