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.	prof. Ing. Michael Šebek, DrSc.	prof. Mgr. Petr Páta, Ph.D.
podpis vedoucí(ho) práce	podpis vedoucí(ho) ústavu/katedry	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

Multirobot, autonomous, BLAH BLAH BLAH

2 MBZIRC Contest

The contest took place in March in Abu Dhabi. Whole competition consisted of three challenges and the grand challenge which connected all challenges together. The first challenge was the only one which was focused solely on UAVs. The goal of the first challenge was to pop multiple big color balloons and to catch small ball carried by the organizer's drone. The other two challenges was designed for both UGVs and UAVs. The second challenge was about building a wall using the robots. Multiple polystyrene bricks were placed in the arena and the robots should have move these bricks to the destination area and stack them on top of each other to build the wall. Lastly the third challenge was to extinguish fire on the surface of the building model. This task was inspired by fires in the high-rise buildings in UAE. UAVs and UGVs carried tanks full of water and squirt it into the fire dummy.

2.1 Second challenge

This thesis is focused on the second challenge and more specifically on the ground robot section of the second challenge, so that we provide more detailed description of this challenge. Each team is given thirty minutes to explore the arena, localize all interest areas and build the wall. There are four types of the bricks with different colors. All bricks must be very light to enable the UAVs to pick them up. The dimension and colors of the bricks can be seen in the figure 1.

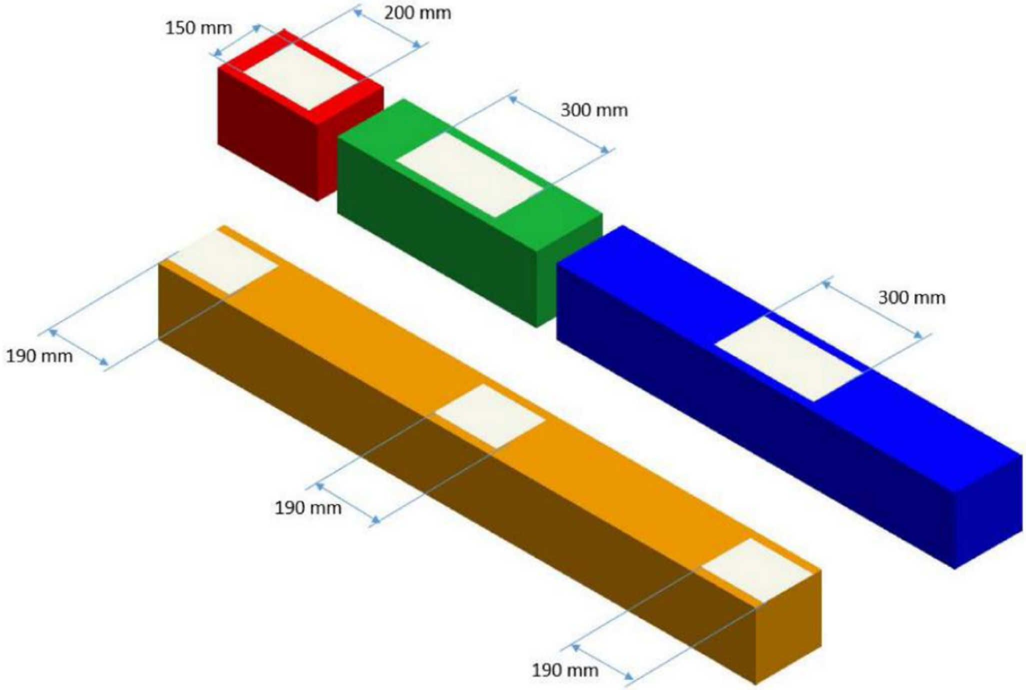

Figure 1: Colors and dimensions of the bricks provided by the organizer.

Each brick has thin metal plate on top of it, so that the robots are able to pick them up using electromagnets. In the beginning of challenge are all bricks placed at beforehand unknown position. Initial position of bricks is unknown but there is predefined pattern in which are the bricks put together. There are different patterns for the UGV piles of bricks and the UAV piles. The UGV bricks are stacked into the multiple height levels whereas the UAV bricks are stacked into the width and all are put on the ground. Due to the low weight of the bricks it is necessary to put UAV bricks into the rails, otherwise the bricks can be easily blown away by the propellers of the drones.

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