Point Cloud Based Reinforcement Learning for Sim-to-Real and Partial Observability in Visual Navigation

Presenter:王鑫怡、邱望弘杰、张郁洁、黄宝仪、倪健翔

2023.3.27

仿生设计与学习实验室
Bionic Design & Learning Lab
In Collaboration with MIT
AncoraSIR.com
SUSTech
Southern University of Science and Technology

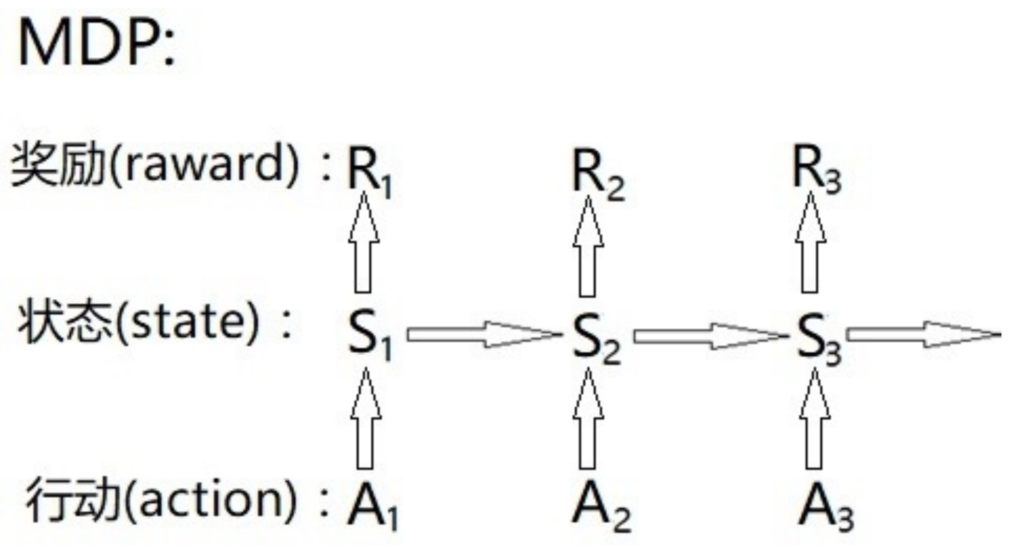
Background

Complex Robotic Tasks

- Locomotion, Navigation , Manipulation, etc
- Having a sequential nature
- Difficult to model

Reinforcement Learning(RL)

- Modeled as Markov Decision Process(MDP)
- Learning a policy by interactions with the environment
- Focus on the desired behavior rather than implementation



Main Problem

Property of RL

- Require an enormous amount of data for training
- Training in the robot is time-consuming and dangerous
- Rely on training in the simulated environment

Simulated Environment

Real Environment

Main Problem

Property of RL

- Observations are used instead of states

Limited Observability Problem

- Observability refers to the ability:
  output data/observation → current state
- Limited observability proper of robotic tasks

Limitations of Current Solutions

- Hard to handle long sequences
- Task-specific, Poor generalization

Related Work

A: Visual Navigation

- Former: perform poorly when used on unseen environments, due to a high number of environment-specific parameters.
- New: using RL agents solve visual navigation tasks have almost perfect SPL score in simulations.

B: Sim-to-Real

- Using photorealistic simulators reconstruct real environments.
- Training agents in environments → A successful Sim-to-Real transfer obtained
- Cons: Simulated robot's performance is strongly affected when it presents changes between training and evaluation.

Proposed Method

The proposed method for sim-to-real transfer focuses on building a canonical space that enables observations to be projected in a standard form independent of the robot and non-essential environmental characteristics.

- The method uses depth sensors, such as RGBD or depth images, but instead of considering their information as images, they are treated as point clouds.
- These point clouds are projected onto a fixed frame of the robot to obtain a collection of points that are agnostic of the camera itself.
- rew(t)=−slackness−distgoal(t)+success(t)

AncoraSIR.com

Theory

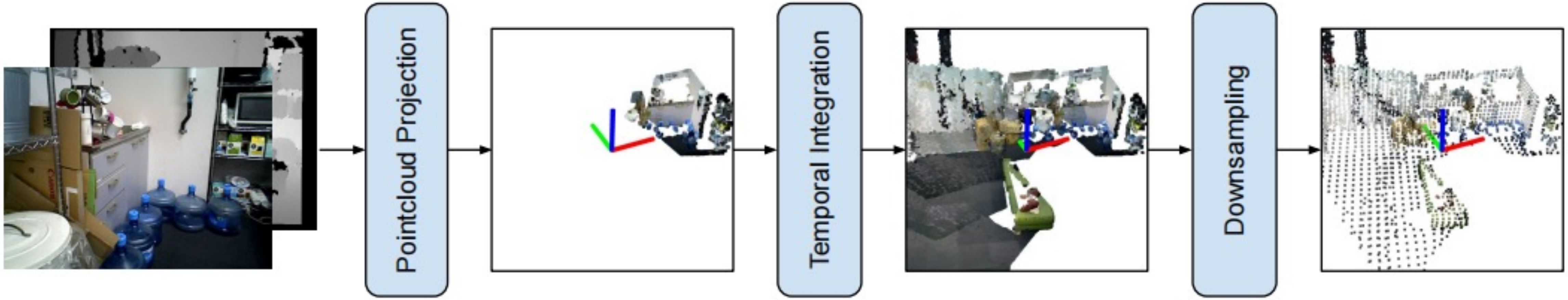
The method based that:

- The use of point clouds is preferred over other 3D representations like meshes and shape primitives because point clouds are directly available from common robotic sensors.

How does this method implement the research:

- The proposed canonical space can overcome the differences between simulations and reality, which can be separated into differences between environments, changes in the observation space, and differences in the action space.

Pipeline of The Proposed Method

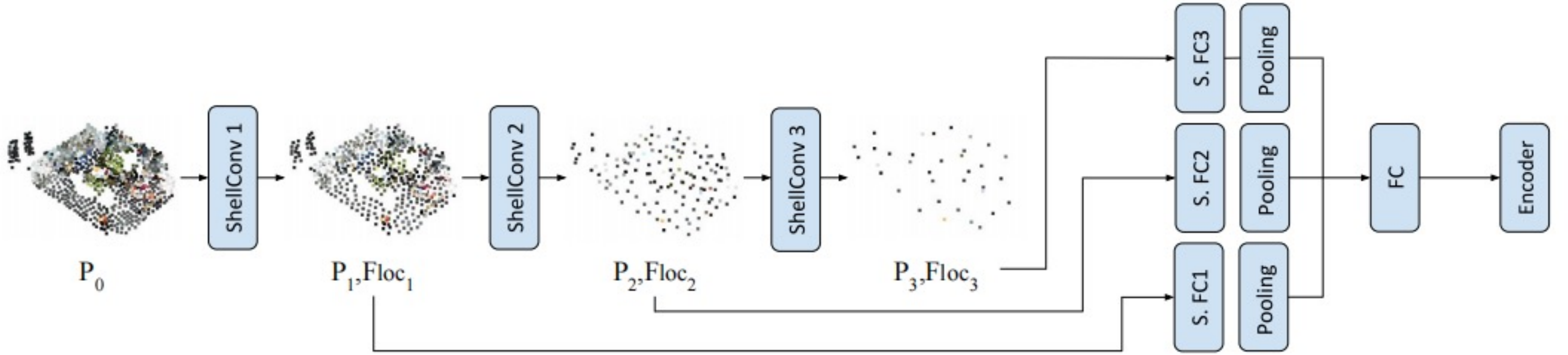


- The RGBD image is projected into a robot's frame (e.g., the robot's base) using robot specific information, making it highly invariant to the particular robot. Afterward, previous observations are explicitly integrated, addressing partial observability. Finally, the resulting point cloud is down-sampled due to practical considerations.

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Modified *ShellNet* Architecture



- *Pi* has dimensions *Ni* ×3, *Floci* have dimensions *Ni* ×*fi* , where *Ni* is the decreasing number of points at each layer, and *fi* is the number of features. Local features do not contain information of their associated points, so that particular information is added using shared-weights fully connected layers, and then pooled to obtain a fifixed-size vector. Finally, features from different levels are combined using similar arguments.

Experimental Result

The experimental results of the paper can be divided into two main parts: sim-to-sim and sim-to-real experiments.

Sim to Sim:

baseline	experiment
RGBD	MP3D and Gibson

Sim to Real: HSR robot， ORB-SLAM2

Set-Up：sim-to-sim

Set-Up

- Two different simulation environments for experiments: MP3D and Gibson, for simulating more challenging and simpler situations. The baseline is RGBD.

Analysis

- Point cloud-based reinforcement learning methods have excellent performance and generalization ability in visual navigation, especially when faced with more challenging situations.

Result

TABLE IV: Sim-to-sim results for the RGBD-based agent

		Train MP3D	Train Gibson
Eval MP3D	Reward	9.21 ± 0.27	2.06 ± 0.35
	SPL	0.44 ± 0.02	0.11 ± 0.02
	Success Rate	0.54 ± 0.02	0.15 ± 0.03
Eval Gibson	Reward	4.44 ± 0.35	4.55 ± 0.21
	SPL	0.27 ± 0.01	0.28 ± 0.01
	Success Rate	0.55 ± 0.01	0.40 ± 0.01

TABLE V: Sim-to-sim results for the point cloud-based agent

		Train MP3D	Train Gibson
Eval MP3D	Reward	13.91 ± 0.16	9.51 ± 0.55
	SPL	0.68 ± 0.02	0.45 ± 0.02
	Success Rate	0.82 ± 0.02	0.64 ± 0.03
Eval Gibson	Reward	9.16 ± 0.33	9.52 ± 0.23
	SPL	0.54 ± 0.01	0.56 ± 0.01
	Success Rate	0.73 ± 0.01	0.74 ± 0.01

Set-Up：sim-to-real

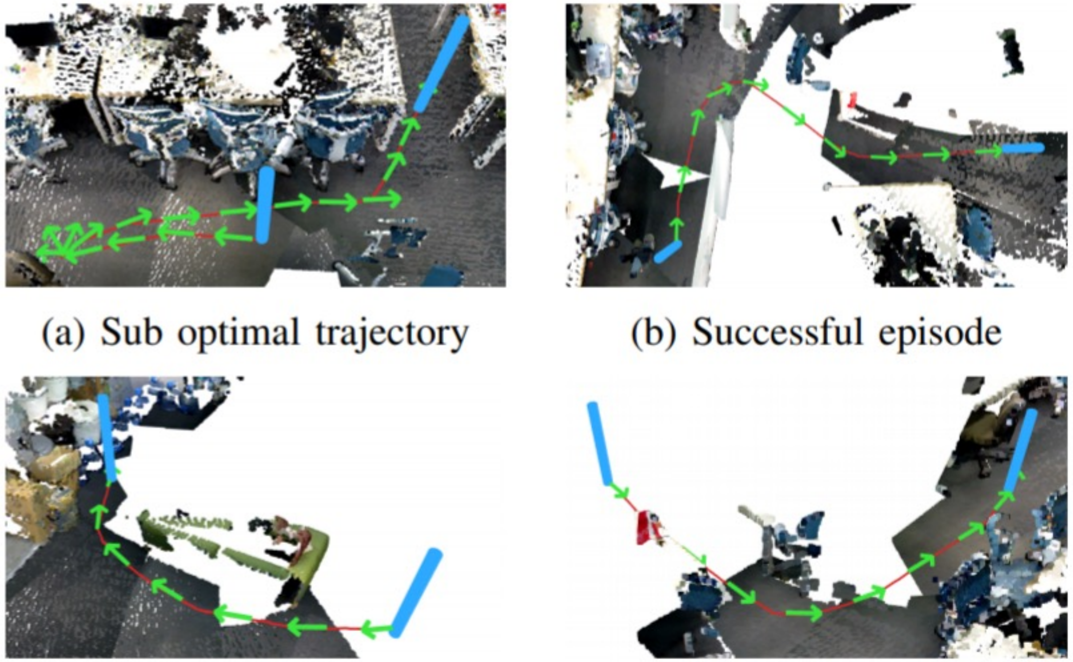
Set-Up

- Use a Toyota HSR robot to conduct experiments to verify the performance.
- Real-world methods require integration of observations through a time-integrated system: ORB-SLAM2.
- Used the full approach in experiments: domain randomization in simulated training, and simulated domain randomization in the real world...

Analysis

- Relative to RGBD-baseline, the proposed agent can navigate through most configurations, solving many non-trivial cases.
- Most suboptimal trajectories occurred when the target was behind the robot, while most failures were collisions with objects that were never in view, and collisions with edges of difficult objects such as chair legs.

Result



(a) Sub optimal trajectory (b) Successful episode

(c) Successful episode (d) Successful episode

AncoraSIR.com

Future Work for Paper

Project extension from the author:

- Most errors corresponds to objects which were never seen, RL signals not be enough

  → Directlyself-supervised curiosity & Common sense in the point cloud space
- Presents more benefits for more challenging visual navigation tasks.

From our points of view:

- Introduce supervised learning in reinforcement learning to provide the correct strategy as supervision information.

Summary

- **Contribution:**
  - (1)Propose a method propose a method to achieve out-of-the-box sim-to-real;
  - (2)Explicitly address the limited observability proper of robotic tasks;
  - (3)Present a point cloud network design that extracts multi-scale features.

Extended Readings

Cited in papers:

[1] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” Nature, vol. 518, no. 7540, pp. 529–533, 2015.

[2] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, “Target-driven visual navigation in indoor scenes using deep reinforcement learning,” in IEEE International Conference on Robotics and Automation (ICRA), 2017.

[3] J. Zhang, L. Tai, P. Yun, Y. Xiong, M. Liu, J. Boedecker, and W. Burgard, “Vr-goggles for robots: Real-to-sim domain adaptation for visual control,” IEEE Robotics and Automation Letters, vol. 4, no. 2, pp. 1148–1155, April 2019.

[4] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, “Sim-to_x0002_real transfer of robotic control with dynamics randomization,” in IEEE International Conference on Robotics and Automation (ICRA), 2018.

[5] M. Jaderberg, V. Mnih, W. M. Czarnecki, T. Schaul, J. Z. Leibo, D. Silver, and K. Kavukcuoglu, “Reinforcement learning with unsupervised auxiliary tasks,” in International Conference on Learning Representations (ICLR), 2017.

[6] F. Leiva, K. Lobos-Tsunekawa, and J. Ruiz-del Solar, “Collision Avoidance for Indoor Service Robots Through Multimodal Deep Reinforcement Learning,” in RoboCup 2019: Robot World Cup XXIII, 2019.

Q&A