Junting Chen

Zurich, Switzerland
Home page | sgt79911030@gmail.com | Junting Chen

EDUCATION

Ph.D. in ISEP, President's Graduate Fellowship — **August 2024 –**
National University of Singapore — *Singapore*

M.Sc. in Robotics, systems, and Control — **September 2020 – December 2023**
ETH Zurich — *Zurich, Switzerland*

B.E. in Computer Science and Technology — **September 2015 – July 2019**
University of Chinese Academy of Sciences — *Beijing, China*
GPA: 3.93/4, Ranking: 2/86

SKILLS

- Robotics
- ROS Development
- Visual Navigation
- Object Manipulation
- Computer Vision
- Large Language Models
- Reinforcement Learning
- Embodied AI

WORK EXPERIENCE

Intern Computer Vision Researcher — **July. 2023 – Now**
Shanghai AI Laboratory — *Shanghai, China*

Intern Perception Engineer — **Jan. 2023 – June. 2023**
ANYbotics AG — *Zurich, Switzerland*

Teaching Assistant, PAI 2021 — **Sep. 2021 – Feb. 2022**
ETH Zurich — *Zurich, Switzerland*

Research Assistant — **Sep. 2019 – May 2020**
Institute of Computing Technology, Chinese Academy of Sciences — *Beijing, China*

PROJECTS/ CHALLENGES

Incremental 3D Scene Graph Construction for high-level planning — Mar. 2021- July. 2021

- Supervisor: Lukas Schmid (schmluk@ethz.ch) Prof. Roland Siegwart (rolandsi@ethz.ch)
- In this **semester project**, I built a ROS package to construct a hierarchical scene graph for objects and rooms in a scene **incrementally** in the runtime, on top of Voxblox++
- Ros package functionality: Predict inter-object semantic relations. Segment free space into different room nodes. Predict room class label based on object types in the room

Robothor Visual Navigation Challenge on CVPR 2020 — Feb. 2020- May. 2020

- Supervisor: Prof. Ruiping Wang (ruiping.wang@vipl.ict.ac.cn)
- This **challenge** requires an agent to navigate to the location of an object specified by its name, with RGBD images stream and current pose as input.
- CNN+LSTM+A2C as the backbone network. Multi-tasking mid-level features are extracted from input RGB images including ResNet features, estimated depth map, segmentation map, estimated surface normal, etc. All features are concatenated together with word embedding and then are fed into a normal actor-critic network to generate actions.
- Result: **Second place** on the leaderboard.

Visiual Navigation Robotic System development — Jun. 2019 – Dec.2019

- Supervisor: Prof. Ruiping Wang (ruiping.wang@vipl.ict.ac.cn)
- Trained an end-to-end object goal navigation neural network by reinforcement learning, which takes only RGBD visual observation as input and predicts the next action. (Python implementation)
- Deployed the object goal navigation model onto the Segway Loomo robot, which runs in the **real world** to search for the target object. (Java implementation)

PUBLICATIONS

[1] J. Chen, G. Li, S. Kumar, B. Ghanem, and F. Yu. "How To Not Train Your Dragon: Training-free Embodied Object Goal Navigation with Semantic Frontiers". In: *Proceedings of Robotics: Science and Systems(RSS)*. 2023. DOI: 10.15607/RSS.2023.XIX.075.

[2] J. Chen, Y. Mu, Q. Yu, T. Wei, S. Wu, Z. Yuan, Z. Liang, C. Yang, K. Zhang, W. Shao, Y. Qiao, H. Xu, M. Ding, and P. Luo. *RoboScript: Code Generation for Free-Form Manipulation Tasks across Real and Simulation*. 2024. arXiv: 2402.14623 [cs.RO].

[3] Y. Mu, J. Chen, Q. Zhang, S. Chen, Q. Yu, C. Ge, R. Chen, Z. Liang, M. Hu, C. Tao, P. Sun, H. Yu, C. Yang, W. Shao, W. Wang, J. Dai, Y. Qiao, M. Ding, and P. Luo. *RoboCodeX: Multimodal Code Generation for Robotic Behavior Synthesis*. 2024. arXiv: 2402.16117 [cs.RO].