

MINZHAO ZHU

Phone: (+86) 156-1166-2982 Email: minzhaozhu@gmail.com

EDUCATION

Beijing Institute of Technology (985/211)

Master of Control Science and Engineering

Sep. 2017 - Jun. 2020

Research Topic: UAV-UGV Collaborative Perception

Supervisor: Mengyin Fu, Yi Yang

Bachelor of Automation

Sep. 2013 - Jun. 2017

GPA: 92.8/100, **Rank:3/172**

WORK EXPERIENCE

ByteDance

Research Engineer at AI Lab

Beijing, China

Jul. 2020 - Present

PUBLICATIONS

[1] **Minzhao Zhu***, Binglei Zhao*, et al, "Navigating to Objects in Unseen Environments by Distance Prediction," arXiv:2202.03735 (IROS2022 Accepted).

[2] Mengyin Fu, **Minzhao Zhu**, et al, "LiDAR-based Vehicle Localization on the Satellite Image via a Neural Network", Robotics and Autonomous Systems, vol.129, 103519, July 2020.

RESEARCHES

Vision-based Object Goal Navigation by Distance Prediction

Jul. 2020 - Mar. 2022

- Object Goal Navigation (ObjectNav) task is to navigate an agent to given target objects in unseen environments. We propose to solve it by predicting the distance to the target on the map.
- A neural network takes an explored semantic map as input, and estimates the distance from map cells to the target object based on learned spatial relations between objects. With such distance map, the agent explores and searches for the target using either classical planning methods or learned policy.
- Our method outperforms the RL-based baseline with only 12% training data.

LiDAR-based Vehicle Localization on the Satellite Image

Oct. 2018 - Jul. 2020

- Propose a vehicle localization method using satellite images in GNSS-denied areas.
- A neural network is proposed to compare the spatial-discriminative feature maps of LiDAR grid-map and satellite image patches. The network outputs the probabilities of correspondence, based on which a particle filter obtains the probability distribution of the vehicle's pose.
- Our method is more robust than baseline methods. The average position error on several KITTI sequences is reduced from 9.64 m to 2.57 m.

PROJECTS

Habitat Challenge 2022 (ObjectNav, HM3D-Semantic)

Apr. 2022 - Aug. 2022

- Propose an ObjectNav method that combines a map-based method and an imitation learning model.
- The map-based method chooses a mid-term goal according to the probability distribution of the target position, which is estimated based on object relations learned from the Visual Genome dataset.
- The IL model is learned from human demonstrations in Habitat-Web dataset.
- A switching strategy selects a final action from one of the two outputs according to the history path, the mid-term goal, and the traversable map.
- First place in the testset as of Sep. 1

Indoor Object Rearrange Robot

Jul. 2020 - Present

- Build a robot equipped with a Kinova arm to perform object search and grasp tasks.
- Develop a Semantic SLAM system fusing 3D LiDAR, camera, IMU, and wheel odometry. The position error is 2.4 cm. The orientation error is 0.74° .
- The aforementioned ObjectNav method is transferred to this robot.

UAV-UGV Collaborative Target Searching System

Jun. 2018 - Jul. 2020

- Lead a team to develop the software of the UAV-UGV collaborative system.
- The UAV constructs a 2D map of an area (3 km \times 3 km) and searches for a moving target (a vehicle). Once the UAV finds the target, it will follow the target and guide the UGV to approach the target.
- To improve LiDAR-based UGV localization in GNSS-denied areas, a neural network estimates the pose offset between the LiDAR grid-map and aerial image patch. The offset is added to the factor graph of LeGO-LOAM to optimize the poses.

HONORS AND AWARDS

2020	National Award for Science and Technology Progress in China	National First Prize
2020	Excellent Postgraduate Thesis of Beijing Institute of Technology	Top 4% in BIT
2018	National Scholarship	Top 2% in BIT
2018	Unmanned System Challenge (Air-Ground Collaboration Contest)	National First Prize
2017	2017 China Autonomous Car Future Challenge	National First Prize

SKILLS

- **GRE:** 324+4.0
- **Programming languages:** C/C++, Python
- **Tools:** ROS, Apollo Cyber RT, OpenCV, Git, PyTorch, Detectron2
- **Familiar Methods:**
SLAM: ORB-SLAM, VINS-Mono, LOAM, LIO-SAM, Fast LIO, R3LIVE
Vision: YOLO, Faster R-CNN, MaskRCNN, CenterNet, ViT, DETR
Point Cloud: PointNet, PointNet++, PointPillars
RL: DQN, A3C, PPO