

# Zhifeng Gu

zhifeng.gu@connect.polyu.hk | <https://neal2020github.github.io>

## Education

|                                                                                                                                                                                             |                     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>The Hong Kong Polytechnic University</b> , Hong Kong                                                                                                                                     | Sep 2023 – Present  |
| <ul style="list-style-type: none"><li>• Doctor of Philosophy, Advised by Prof. Bing Wang</li><li>• <b>Research Interests:</b> Embodied Agents, Spatial Intelligence, World Models</li></ul> |                     |
| <b>Beihang University</b> , Beijing, China                                                                                                                                                  | Sep 2019 – Jun 2023 |
| <ul style="list-style-type: none"><li>• Bachelor of Software Engineering, Advised by Prof. Lu Sheng</li></ul>                                                                               | <b>GPA:</b> 3.8/4.0 |

## Internships

|                                                                                                                                  |                    |
|----------------------------------------------------------------------------------------------------------------------------------|--------------------|
| <b>XPENG Robotics</b> , Shenzhen, China                                                                                          | Apr 2026 – Present |
| <ul style="list-style-type: none"><li>• <b>Research Direction:</b> Recursive Self-Improvement in Embodied Intelligence</li></ul> |                    |

## Publications

### R<sup>3</sup>L: Reasoning 3D Layouts from Relative Spatial Relations

Zhifeng Gu\*, Yuqi Wang\*, Bing Wang

ICML 2026

### MMOne: Representing Multiple Modalities in One Scene

Zhifeng Gu, Bing Wang

ICCV 2025

## Research Experience

|                                                                                                                                                                                                                                                                                                                                                                            |                     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------|
| <b>Spatial Reasoning for 3D Layout Generation</b>                                                                                                                                                                                                                                                                                                                          | Jul 2025 – Jan 2026 |
| SPACE Group, HK PolyU, Advised by Prof. Bing Wang                                                                                                                                                                                                                                                                                                                          |                     |
| <ul style="list-style-type: none"><li>• Present a general framework for 3D layout generation, featuring invariant spatial decomposition and consistent spatial imagination to improve the reliability and consistency of multi-hop relative spatial reasoning.</li></ul>                                                                                                   |                     |
| <b>Multimodal Scene Representation</b>                                                                                                                                                                                                                                                                                                                                     | Sep 2024 – Mar 2025 |
| SPACE Group, HK PolyU, Advised by Prof. Bing Wang                                                                                                                                                                                                                                                                                                                          |                     |
| <ul style="list-style-type: none"><li>• Propose a general framework for representing multiple modalities in one scene, with a modality modeling module and a multimodal decomposition mechanism to address modality conflicts.</li></ul>                                                                                                                                   |                     |
| <b>6D Object Pose Estimation and Tracking</b>                                                                                                                                                                                                                                                                                                                              | Jun 2022 – Jun 2023 |
| VCG Lab, Beihang University, Advised by Prof. Lu Sheng                                                                                                                                                                                                                                                                                                                     |                     |
| <ul style="list-style-type: none"><li>• Study 6D object pose tracking in human-object interaction scenes, where hand occlusion poses great challenges.</li><li>• Develop a motion-centric framework that explicitly models object motion across successive frames and refines pose estimation through shape completion and attention-based temporal integration.</li></ul> |                     |

## Selected Honors and Awards

- Outstanding Graduate of Beihang University

## Selected Skills

Python, PyTorch, C/C++, CUDA, Blender, etc.