Qihang Peng

+86 187 3873 2400 | pqh22@mails.tsinghua.edu.cn | LinkedIn | GitHub | Beijing, China

Education

Tsinghua University Beijing, China

Undergraduate student of Mechanics and Vehicle engineering; ***GPA: 3.90/4.00*** *Sep 2022 – Present*

Core Course: Mathematical Analysis(4.0); Advanced Algebra(4.0); Probability and Statistics(4.0); Tensor Analysis(4.0); ODE(4.0); PDE(4.0); Theoretical Mechanics(4.0); Programming Fundamentals(4.0); Pattern Recognition and Machine Learning(4.0)

Skills & Interests

Skills: C/C++ and Python programming; Pytorch framework

Interests: 3D Vision; Foundation Models; Embodied Intelligence

Awards

National Scholarship: Highest honor for undergraduates in China (Oct 2024)

Academic Excellence Award of Tsinghua University: Top GPA in Tsinghua University (Oct 2024)

Outstanding Championship and Innovation Award in AGC2024: 1st place in the Track on Multi-View 3D Visual Grounding of the Autonomous Grand Chanllenge at CVPR 2024 (Jun 2024)

Wang Dazhong Scholarship: Best comprehensive excellence scholarship in Tsinghua University. Only one in each department every year.(Oct 2023)

Publications

Qihang Peng, Henry Zheng, Gao Huang. ProxyTransformation: Preshaping Point Cloud Manifold With Proxy Attention For 3D Visual Grounding. Submitted to *IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2025*

Henry Zheng*, Hao Shi*, **Qihang Peng**, Yong Xien Chng, Rui Huang, Yepeng Weng, zhongchao shi, Gao Huang. DenseGrounding: Improving Dense Language-Vision Semantics for Ego-centric 3D Visual Grounding. *International Conference on Learning Representations (**ICLR**), 2025*

Selected Projects

ProxyTransformation: Preshaping Point Cloud Manifold With Proxy Attention For Visual Grounding

- Guided by Gao Huang, Department of Automation, Tsinghua University
- We propose **Proxy Transformation** suitable for multimodal task to efficiently improve the point cloud manifold. Our method first leverages **Deformable Point Clustering** to identify the point cloud sub-manifolds in target regions. Then, we propose a **Proxy Attention** module that utilizes multimodal proxies to guide point cloud transformation. Built upon Proxy Attention, we design a submanifold transformation generation module where textual information globally guides translation vectors for different submanifolds, optimizing relative spatial relationships of target regions. Simultaneously, image information guides linear transformations within each submanifold, refining the local point cloud manifold of target regions. (Oct 2024 – Present)
- Supported by the **Beijing Natural Science Foundation** Undergraduate Research Program.

DenseGrounding: Improving Dense Language-Vision Semantics for Ego-centric 3D Visual Grounding

- Guided by Gao Huang, Department of Automation, Tsinghua University
- For visual features, we introduce the Hierarchical Scene Semantic Enhancer, which retains dense semantics by capturing fine-grained global scene features and facilitating cross-modal alignment. For text descriptions, we propose a Language Semantic Enhancer that leverage large language models to provide rich context and diverse language descriptions with additional context during model training. (Feb 2024 – June 2024)
- **Outstanding Championship** and **Innovation Award** in the Track on Multi-View 3D Visual Grounding of the Autonomous Grand Chanllenge at CVPR 2024.

Robust multimodal object detection with 4D mmwave radar

- Guided by Shaobing Xu, School of Vehicle and Mobility, Tsinghua University

- To address the challenges of low signal-to-noise ratio and difficult feature extraction in 4D millimeter-wave radar data, we replicated related works and optimized the neural network architecture. We also attempted noise reduction and data densification while deepening our understanding of 4D millimeter-wave radar. (Apr 2023 – Dec 2023)
- Supported by Tsinghua University Academic Advancement Program and I'm the leader.