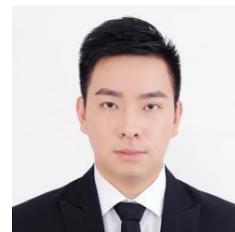


Liu Zihua Tokyo, Japan

TEL: 09050024487 | E-Mail: liu.z.aj@m.titech.ac.jp | Github: github.com/Magicboomliu

HomePage: <https://magicboomliu.github.io/>

Google Scholar: <https://scholar.google.com/citations?user=5GVxjvIAAAAJ&hl=en>



Education

Undergraduate: South China University of Technology (SCUT) 985&211 2016/09-2020/06

Major: Information and Interactive design **Bachelor of Engineering** | GPA 3.78/4.00

Master: Institute of Science Tokyo (Tokyo Institute of Technology) 2021/04-2023/03

Major: System Control and Engineering Okutomi & Monno Lab **Master of Engineering** | GPA 4.00/4.50 (Top5%)

PhD Student: Institute of Science Tokyo (Tokyo Institute of Technology) 2023/03-2026/03

Major: System Control and Engineering Okutomi & Monno Lab **Doctor of Engineering**

Selected Publications

1. *Digging Into Normal Incorporated Stereo Matching*, ACM International Conference on Multimedia (ACMMM 2022).

Zihua Liu, Songyan Zhang, Zicheng Wang and Masatoshi Okutomi.

2. *Global Occlusion-Aware Transformer for Robust Stereo Matching*. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV 2024)

Zihua Liu, Yizhou Li and Masatoshi Okutomi.

3. CFDNet: A Generalizable Foggy Stereo Matching Network with Contrastive Feature Distillation. International Conference on Robotics and Automation (ICRA2024)

Zihua Liu, Yizhou Li and Masatoshi Okutomi.

4. VSRD: Volumetric Silhouette Rendering for Weakly Supervised 3D Object Detection. Proceedings of the IEEE/CVF Computer Vision and Pattern Recognition. (CVPR 2024)

Zihua Liu*, Hiroki Sakuma*, Masatoshi Okutomi (Equal Contribution)

Internship Experience

1. Sony Semiconductor Solutions Tokyo 2025/01~2025/06

Position: R&D Research Internship.

Using multi-modal information like LiDAR, images, bounding boxes and textual prompts to training a multi-view diffusion models for street-view and BEV map generations for autonomous preception system.

2. Preferred Networks. Tokyo 2024/08~2024/10

Position: R&D Research Internship.

I conducted research on background generation using diffusion models, focusing on generating foregrounds in various artistic styles. The proposed method allows for training at a lower cost compared to mainstream approaches such as LoRA and ControlNet.

3. Sensetime Japan. Tokyo 2023/06~2023/11

Position: Computer Vision Internship.

As a research intern, I co-developed VSRD, a method for monocular 3D object detection with weak 2D supervision, bypassing 3D labels. Our approach generates pseudo labels via multi-view auto-labeling and uses Instance-aware Volumetric Silhouette Rendering to optimize 3D bounding boxes without relying on labelled 3D data or LiDAR point cloud.

4. Megvii Shanghai Research Institute. Shanghai 2022/02~2022/04

Position: Computer Vision Internship.

(1) Focused on parking slot detection and segmentation for autonomous driving. Introduced an angle constraint for parking slot corners, boosting detection recall by 10%.

(2) Developed a novel data augmentation strategy to maximize limited training data, significantly improving detection in challenging areas such as image edges and occluded regions. The solution was adopted by Megvii as part of their BEV pipeline for automatic parking algorithms.

Language Skills

English: CET-6 580 | TOFEL 91 Japanese: JPLT N1 131

Academic Services

I serve as a reviewer for top international conferences including ICLR, NeruIPS, CVPR, ICML, etc.

Professional Skills.

- Proficient in Python scientific programming, have a good command of using Pytorch for model design and reproduction.
- Proficient in computer vision algorithms both based on multi-view geometry and deep learning, familiar with and

reproduce the CNNs and transformer based methods. Have a deep understanding of the research and development of 3D vision task algorithms(including depth/flow estimation, SFM/MVS etc).

- Familiar with the diffusion models including stable diffusion and diffusion-based 3D reconstruction.
- Familiar with the differential rendering techniques such as NeRF, 3DGS and diffusion-based novel view synthesis.
- Familiar with common autonomous driving framework such MMDetection and MMDetection3D.
- Understand the basis of the CUDA and C++ based pytorch-extension programing.
- Familiar with linux, git, accelrate, docker operation and cluster computation.
- Have strong mathematical analysis and paper writing skills, good at communication and team works.