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 & Awards

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 Avg.Grade 90.24 / 100

South China University of Technology Scholarship 2019 /12

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

Major: System Control and Engineering Okutomi & Tanaka Computer Vision and Image processing Lab

Master of Engineering | GPA 4.00/4.50 (Top5%) Avg.Grade 95.6/100

SCIE Award(The Society of Instrument and Control Engineering). 2023/03

SPRING Development of Next-Generation Front-Runners with Comprehensive Knowledge and Humanity 2023~2026

Ph.D candiates: Institute of Science Tokyo (Tokyo Institute of Technology) **2023/03-2026/03**

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

Position: R&D Research Internship.

（1） During my time at PFN, 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.

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

Position: Computer Vision Internship.

(1) 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. This work has been accepted at CVPR 2024.

3. Megvii Shanghai Research Institue. 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 IPM pipeline for automatic parking algorithms.

4.PixelShift.AI. Shanghai 2020/10-2021/02

Position: Algorithm Engineer Internship.

(1) Developed an Android mobile and web AR solution for art toys using Google Mediapipe. Contributed to ZOO AR camera R&D, improving AR tracking by replacing ORB and SIFT with Google KNIFT key point matching.

(2) Designed and researched the lip sync algorithm for mobile AR virtual characters (VRM model). Developed a lightweight lip sync solution using MFCC features with shallow MLP and SVM, achieving accurate lip motion data feedback.

Language Skills

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

Academic Services

I serve as a reviewer for International conferences including ICLR, NeruIPS, ACMMM, BMVC, etc.

Professional Skills.

· Proficient in Python scientific programming, with expertise in PyTorch for model design and reproduction.

· Skilled in computer vision algorithms based on multi-view geometry and deep learning, with experience reproducing CNN and Transformer-based methods. Possess in-depth knowledge of 3D vision tasks, including Stereo Matching, Optical Flow Estimation, and Structure from Motion (SfM).

· Experienced with diffusion models, including DDPM and Stable Diffusion.

· Knowledgeable in differential rendering techniques such as NeRF, 3D Gaussian Splatting (3DGS), and diffusion-based novel view synthesis.

· Strong understanding of CUDA and C++-based PyTorch extension programming.

· Proficient in Linux, Git, Accelerate, and Docker operations.

· Excellent mathematical analysis and paper writing skills, with strong communication abilities and a collaborative approach to research and teamwork.