

Jiaming Sun

Homepage: jiamingsun.ml

+86 18503067669 (WeChat)

suenjiaming@gmail.com

[JiamingSuen](#)

[JiamingSuen](#)

suenjiaming@hotmail.com

zoom 253 972 8223

Education

- 2013–2017 **B.Eng in Computer Science and Engineering, SUSTech.**
Thesis title: Agile Flight Control for Quadcopters. Best thesis award of 2017.

Work Experiences

- 2018.11–Now **Computer Vision Researcher, SenseTime.**
Leading an R&D team of 10+ researchers and engineers, working on 3D Computer Vision and its applications in Augmented/Mixed Reality. Affiliated with the MIG-3D&MR Group located in Hangzhou, China. Supervised by Prof. Xiaowei Zhou.
- 2018.6–Now **Research Associate, 3D Vision Group at Zhejiang University.**
Managing undergraduate and junior graduate students' research and student recruiting. Affiliated with the State Key Lab of CAD & CG. Hosted by Prof. Xiaowei Zhou.
- 2017.6–2018.6 **Research Intern, DJI.**
Worked on multiple research projects in an internal autonomous driving initiative now publicly known as DJI Automotive. Advised by Dr. Xiaozhi Chen.

Selected Publications

- CVPR 2021 **NeuralRecon: Real-time Coherent 3D Reconstruction with Monocular Video.**
(Oral) Proposed the first learning-based pipeline that reconstructs 3D scene geometry from a monocular video in real-time. Unlike previous methods that estimate single-view depth maps and perform TSDF fusion later, NeuralRecon jointly reconstructs and fuses local surfaces directly in the sparse volumetric TSDF representation. This design allows the network to capture local smoothness prior and global shape prior of 3D surfaces, resulting in accurate and coherent surface reconstruction. NeuralRecon generalizes well to new data domains and is able to reconstruct large-scale 3D scenes on a laptop GPU in 33 keyframes per second.
- CVPR 2021 **LoFTR: Detector-Free Local Feature Matching with Transformers.**
Proposed a Transformer-based local feature matching pipeline that extracts dense matches at a coarse level first and refines the good matches to a sub-pixel level later on. This design avoids the repeatability bottleneck of interest point detectors that operate only on a single image. Benefiting from the detector-free design and the inductive bias of Transformers (namely global receptive field and positional encoding), LoFTR is capable of extracting high-quality semi-dense matches even in indistinctive regions with low-textures, motion blur, or repetitive patterns.
- CVPR 2020, TPAMI 2021 **Disp R-CNN: Stereo 3D Object Detection via Shape Prior Guided Instance Disparity Estimation.**
Proposed a 3D object detection pipeline that estimates instance-level disparity maps on individual objects. This design guides the network to learn the category-level object shape prior for better disparity estimation and 3D object detection. Even when LiDAR ground-truth is not available at training time, Disp R-CNN outperforms previous state-of-the-art methods by 20% in terms of average precision.

Engineering Projects

- 2019.8 – Now **Learning-based Local Feature Extraction and Matching**, *SenseTime*.
Designed and developed a learning-based local feature extraction and matching pipeline named SVCNN internally. SVCNN has landed in 10+ products in the company's AR platform SenseMARS and deployed to popular mobile platforms, including CPU/GPU/DSP/NPU on Android and iOS. SVCNN outperforms the open-source counterpart SuperPoint by ~20% in terms of success rate in visual localization while runs at ~20FPS on modern smartphone CPUs.
- 2019.2 – Now **Object 6D Pose Estimation and Tracking for AR**, *SenseTime*.
Working on object pose estimation pipelines specialized for AR applications. Developing an integrated object pose estimation and tracking pipeline with a focus on leveraging spatial-temporal memory to improve pose tracking accuracy. Designed an AR-based object pose annotation app that can annotate and generate massive ground-truth poses for a new object in a few minutes.
- 2018.11 – Now **3D Semantic Modeling**, *SenseTime*.
3D panoptic segmentation module in SenseMARS. Designed and developed a web-based semi-automatic 3D semantic label annotation pipeline that can speed up the traditional 2D annotation pipeline (drawing polygon masks) by 20×.
- 2018.1 – 2018.9 **Self-supervised scene flow estimation with CNN**, *DJI & ZJU*.
Iterative cost volume processing for optical flow, disparity and MVS. Synthetic data collection in GTA. Worked on resolving dynamic object motion with a 3D consistency loss.
- 2017 – 2018 **Integrated 3D object detection and tracking**, *DJI*.
Investigated state-of-the-art 3D detectors at the time, including MV3D, AVOD and F-PointNet, etc. Implemented a joint detection and tracking CNN into the internal 3D detection pipeline used on the testing vehicles.
- 2014 – 2017 **Various self-practicing projects in CV and Robotics**, *SUSTech*.
Built a quadcopter testbed from scratch for my undergraduate thesis, including the mechanical structure, propulsion system, embedded flight controller, perception system, data telemetry and motion capture system.

Experiences Summary

- Research Four years (since 2017) of research experience in 3D Computer Vision, for topics including 3D semantics, 3D representation learning, 2D/3D correspondence estimation, shape reconstruction and pose estimation. Rich mentoring experiences, mentored 12+ junior researchers at SenseTime and ZJU.
- Engineering Full-stack 3D Vision and Deep Learning engineering experiences in data acquisition/annotation pipelines, algorithm design and real-world deployment. Rich communication, collaboration and project management skills.

Skills

- Programming Python, C++, CUDA, JavaScript/HTML/CSS, \LaTeX , Mathematica
Library PyTorch, TensorFlow, MXNet, Caffe, ROS
Designing Blender, Adobe CC Suite, FCPX, SolidWorks

Service

- Conference reviewer: CVPR, ICCV, ACCV

Publication List

* denotes equal contribution.

Conference Papers (Peer-reviewed)

- [1] **Jiaming Sun***, Yiming Xie*, Linghao Chen, Xiaowei Zhou, and Hujun Bao. NeuralRecon: Real-Time Coherent 3D Reconstruction from Monocular Video. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. **Oral presentation** and **Best paper candidate**.
- [2] **Jiaming Sun***, Zehong Shen*, Yuang Wang*, Hujun Bao, and Xiaowei Zhou. LoFTR: Detector-Free Local Feature Matching with Transformers. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [3] **Jiaming Sun***, Linghao Chen*, Yiming Xie, Siyu Zhang, Qinhong Jiang, Xiaowei Zhou, and Hujun Bao. Disp R-CNN: Stereo 3D Object Detection via Shape Prior Guided Instance Disparity Estimation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [4] Jianan Zhen*, Qi Fang*, **Jiaming Sun**, Wentao Liu, Wei Jiang, Hujun Bao, and Xiaowei Zhou. SMAP: Single-Shot Multi-Person Absolute 3D Pose Estimation. *European Conference on Computer Vision (ECCV)*, 2020.

Journal Papers

- [5] Linghao Chen*, **Jiaming Sun***, Yiming Xie, Siyu Zhang, Qing Shuai, Qinhong Jiang, Guofeng Zhang, Hujun Bao, and Xiaowei Zhou. Shape Prior Guided Instance Disparity Estimation for 3D Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2021.

References

- Xiaowei Zhou, xwzhou@zju.edu.cn
- Qixing Huang, huangqx@cs.utexas.edu
- Xiaozhi Chen, cxz.thu@gmail.com