XINHAI CHANG

Beijing, China | changxinhai@stu.pku.edu.cn | June 18, 2026

chang-xinhai.github.io | chang-xinhai | Google Scholar

Education

Peking University, Beijing, China — Expected: Jun 2027
Bachelor of Science in Data Science and Artificial Intelligence, Yuanpei College

- GPA: 3.748/4.0
- Coursework: Computer Vision, Machine Learning, Multimodal Learning, Intelligent Robotics [3.9+]

Research Interests

My research bridges **Computer Vision** and **Robotics** to close the loop of **Real-Sim-Real** for Embodied AI. I focus on vision-centric **High-Fidelity 3D Reconstruction and Generation** (Real-to-Sim) to digitize complex physical worlds into simulation-ready assets, and leverage these assets for robotics-centric **Generalizable Robot Policy Learning** (Sim-to-Real) to deploy robust intelligence in the real world.

Technical Skills

- **Robotics and Simulation:** Isaac Lab, Isaac Sim, Curobo, Robot Manipulation, Motion Planning.
- **Languages and Frameworks:** Python, C/C++, Bash, PyTorch, TensorFlow, Pandas, Qt.
- **Algorithms and Vision:** 3D Reconstruction (Gaussian Splatting), Multimodal Learning, 3D Generation.

Research Experience

VCL Lab, Peking University — Jun 2026 - Present
Research Assistant (Advised by Prof. Wenzheng Chen)

- Focused on the full **Real-Sim-Real** pipeline, including world modeling, robot manipulation, and 3D perception.

CoRe Lab, Peking University — Mar 2025 - Jun 2026
Research Assistant (Advised by Prof. Yixin Zhu)

- Focusing on **Sim-to-Real** intelligence, specifically large-scale simulation and robust mobile manipulation.

Media Lab and EECS, MIT — Mar 2024 - Present
Research Assistant (Mentored by Kaichen Zhou)

- Focusing on **Real-to-Sim** pipelines, working on high-fidelity 3D reconstruction and digital twin creation.

Publications

- Niu, Y.*, **Chang, X.***, Liu, X.*, Jiao, Z., Zhu, Y. (2026). *A Scalable Whole-body Trajectory Generator for Coordinated Mobile Manipulation*. The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) **Highlight**.
- Zhou, K., Chen, Y., Zhan, F., Hua, H., Chen, G., **Chang, X.**, Qu, A., Du, Y., Liu, Z., Liang, P. P., Wang, M. (2026). *GEM-4D: Geometry-Enhanced Video World Models for Robot Manipulation*. European Conference on Computer Vision (ECCV).
- Li, H.*, Huang, X.*, **Chang, X.***, Zhou, J., Zhao, H. (2025). *TerraX: Visual Terrain Classification Enhanced by Vision-Language Models*. International Conference on Intelligent Robots and Systems (IROS) (Oral).
- Zhou, K., Wang, Y.*, Chen, G.*, **Chang, X.***, Beaudouin, G., Zhan, F., Liang, P. P., Wang, M. (2026). *PAGE-4D: Disentangled Pose and Geometry Estimation for 4D Perception*. International Conference on Learning Representations (ICLR).
- **Chang, X.**, Zhou, K. (2026). *Neural Surface Reconstruction from Sparse Views Using Epipolar Geometry*. CVPR 2026 Workshop A2AMML.
- Zhou, K.*, Hong, L.*, **Chang, X.***, Zhong, Y., Xie, E., Dong, H., Li, Z., Yang, Y., Li, Z., Zhang, W. (2025). *SplatMesh: Interactive 3D Segmentation and Editing Using Mesh-Based Gaussian Splatting*. CVPR Workshop (CVPRW).

Preprints

- Zhou, K.*, **Chang, X.***, Kim, T.*, Zhang, J.*, Cao, Y., Peng, C., Zhan, F., Zhao, H., Dong, H., Ting, K. M., Zhu, Y. (2026). *RAD: A Realistic Multi-View Benchmark for Pose-Agnostic Anomaly Detection*. Under Review.
- Zhou, K.*, Bai, Z.*, **Chang, X.***, Wang, M., Liang, P. P., Zhan, F. (2026). *Stream3D: Sequential Multi-View 3D Generation via Evidential Memory*. Preprint.
- Li, J., Zhang, B., **Chang, X.**, Chen, W. (2026). *NSL-SLAM: High-Fidelity Neural Structured-Light Depth for Practical SLAM and Reconstruction*. Preprint.

* denotes equal contribution

Academic Projects

● *Computer Vision* ● *Robotics*

I. Real-to-Sim: High-Fidelity 3D Reconstruction and Generation

EpiS: Neural Surface Reconstruction from Sparse Views Using Epipolar Geometry CVPRW 2026
- A neural surface reconstruction model resolving geometric ambiguity via epipolar geometry and depth priors.
- Responsible for epipolar feature extraction, monocular depth alignment, and baseline evaluation.
- Roadmap: Achieves high-fidelity 3D geometry reconstruction from sparse visual observations. ●

SplatMesh: Interactive 3D Segmentation and Editing CVPRW 2025
- An interactive 3D segmentation and editing algorithm integrating Gaussian Splatting with QEM downsampling.
- Responsible for QEM downsampling, mesh editing pipeline construction, and model training and evaluation.
- Roadmap: Generates editable, memory-efficient digital twins for downstream simulation tasks. ●

Stream3D: Sequential Multi-View 3D Generation via Evidential Memory Preprint
- A sequential multi-view 3D generation method using evidential memory for consistent high-quality generation.
- Roadmap: Enables consistent 3D asset generation for scalable digital twin creation. ●

PAGE-4D: Disentangled Pose and Geometry Estimation for 4D Perception ICLR 2026
- A feedforward 4D perception model featuring a dynamics-aware mask to disentangle static and dynamic scenes.
- Responsible for novel view synthesis module, LVSM architecture adaptation, and baseline evaluation.
- Roadmap: Extends high-fidelity reconstruction to dynamic environments for 4D simulation. ●

RAD: Realistic Benchmark for Pose-Agnostic Anomaly Detection Preprint
- A realistic, robot-captured multi-view benchmark for anomaly detection on reflective and symmetric objects.
- Responsible for dataset curation, 2D and 3D model training, and baseline evaluation.
- Roadmap: Benchmarks high-fidelity perception against unstructured physical complexities. ● ●

II. Sim-to-Real: Generalizable Robot Policy Learning

AutoMoMa: Scalable Coordinated Mobile Manipulation Trajectory Generation CVPR 2026 Highlight
- A scalable data generation pipeline utilizing GPU-accelerated planning for coordinated mobile manipulation.
- Responsible for scene generation, large-scale simulation data collection, and policy training and evaluation.
- Roadmap: Generates large-scale manipulation data to enable generalizable policy learning. ●

TerraX: Visual Terrain Classification Enhanced by VLMs IROS 2025
- A vision-language framework combining multimodal data and a CLIP-based VLM for terrain classification.
- Responsible for heterogeneous data organization, fine-grained annotation, and baseline evaluation.
- Roadmap: Enables robust visual terrain perception for generalizable robotic deployment. ● ●

Honors and Awards

- **2025 National Scholarship** (Highest honor for undergraduate students in China)
- 2024 Second-class Scholarship, Peking University
- First Prize, 21st Jiang Zehan Cup Mathematical Modeling Competition, 2024
- Xing Zhengde Scholarship, Yuanpei College, 2023
- Third-class Scholarship, Peking University, 2023

Teaching Experience

Teaching Assistant, Introduction to Computation B (Prof. Jun Sun) Autumn 2025

- Led weekly programming labs, graded assignments, and designed exam questions.

Teaching Assistant, Introduction to Computation C (Prof. Baobao Chang) Autumn 2025

- Led weekly programming labs, graded assignments, and designed exam questions and the final project.

Teaching Assistant, Introduction to Computer Systems (Prof. Lu Zhang) Autumn 2024

- Led weekly discussion sessions, graded assignments, and contributed to exam design.

Extracurricular Activities

- Volunteer Teacher, International Asian Liver Center, Gansu, China, Summer 2024
- Volunteer, Yunnan Golden Monkey Conservation, Peking University, Winter 2023