Gen Li

genli@cs.stanford.edu | Homepage | Google Scholar | LinkedIn | GitHub

Summary

Ph.D. candidate in the direct doctorate program at the Computer Vision and Learning Group (VLG), ETH Zürich, supported by a Microsoft Spatial AI Lab research grant. My research bridges multimodal foundation models, egocentric vision, and robotics toward building scalable foundation agents that understand and interact with complex 3D/4D environments.

Education

ETH Zürich — Zürich, Switzerland
Ph.D. in Computer Science (Direct Doctorate) — *Sep 2020 – Now*
Supervisors: Prof. Marc Pollefeys and Prof. Siyu Tang

Tsinghua University — Beijing, China
B.S. in Computer Science — *Sep 2016 – Jun 2020*

Industry & Research Experience

Meta — *May 2026 – Nov 2026*
Research Scientist Intern
Mentored by Javier Romero.

Google — *Oct 2025 – Jan 2026*
Student Researcher
Mentored by Vassilis Choutas in Thabo Beeler's group. Architected large-scale models and scaled pretraining for temporal 3D human motion reconstruction from monocular RGB video.

Stanford University — *Jun 2019 – Feb 2020*
Undergraduate Visiting Research Intern
Supervised by Prof. Jeannette Bohg. Collaborated on learning topological motion primitives for knot planning (published at IROS 2020).

Publications

[1] **ADAPT: Agile Diffusion Action Priors for Robust and Steerable Online Text-Driven Humanoid Control**
Yan Wu, Chenhao Li, Kaifeng Zhao*, **Gen Li***, Marco Hutter, Siyu Tang.
In submission to *Conference on Robot Learning (CoRL)*, 2026. (* equal contribution)

[2] **ReViV: Reconstructing the Viewer and the View in 4D from Monocular Egocentric Video**
Xiaozhong Lyu*, **Gen Li***, Zhiyin Qian, Xucong Zhang, Marc Pollefeys, Siyu Tang.
European Conference on Computer Vision (ECCV), 2026. (* equal contribution)

[3] **EgoM2P: Egocentric Multimodal Multitask Pretraining**
Gen Li, Yutong Chen, Yiqian Wu, Kaifeng Zhao, Marc Pollefeys, Siyu Tang.
IEEE/CVF International Conference on Computer Vision (ICCV), 2025. **(🏆 EgoVis Distinguished Paper Award)**

[4] **VolumetricSMPL: A Neural Volumetric Body Model for Efficient Interactions, Contacts, and Collisions**

Marko Mihajlovic, Siwei Zhang, **Gen Li**, Kaifeng Zhao, Lea Müller, Siyu Tang.
IEEE/CVF International Conference on Computer Vision (ICCV), 2025. **(Highlight, Top 2.3%)**

[5] **DartControl: A Diffusion-Based Autoregressive Motion Model for Real-Time Text-Driven Motion Control**
Kaifeng Zhao, **Gen Li**, Siyu Tang.
International Conference on Learning Representations (ICLR), 2025. **(Spotlight, Top 3.26%)**

[6] **EgoGen: An Egocentric Synthetic Data Generator**
Gen Li, Kaifeng Zhao, Siwei Zhang, Xiaozhong Lyu, Mihai Dusmanu, Yan Zhang, Marc Pollefeys, Siyu Tang.
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024. **(Oral, Top 0.78%; 🏆 EgoVis Distinguished Paper Award)**

[7] **TRTM: Template-based Reconstruction and Target-oriented Manipulation of Crumpled Cloths**
Wenbo Wang, **Gen Li**, Miguel Zamora, Stelian Coros.
IEEE International Conference on Robotics and Automation (ICRA), 2024.

[8] **Reconstructing Action-Conditioned Human-Object Interactions Using Commonsense Knowledge Priors**
Xi Wang*, **Gen Li***, Yen-Ling Kuo, Muhammed Kocabas, Emre Aksan, Otmar Hilliges.
International Conference on 3D Vision (3DV), 2022. (* equal contribution)

[9] **Learning Topological Motion Primitives for Knot Planning**
Mengyuan Yan, **Gen Li**, Yilin Zhu, Jeannette Bohg.
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.

Invited Talks

- **ICCV 2025:** Invited speaker at the 2nd EgoMotion Workshop. *Oct 2025*
- **Microsoft:** *Toward Egocentric Multimodal Multitask Pretraining* (EgoGen + EgoM2P). *Sep 2025*

Honors & Awards

- EgoVis Distinguished Paper Award (×2) *2026*
  Awarded for papers EgoGen and EgoM2P (Sponsored by Google DeepMind)
- ETH Zürich Direct Doctorate Scholarship *2020*
- Outstanding Bachelor's Graduate, Tsinghua University *2020*
- POSCO Asia Fellowship *2017, 2018, 2019*

Teaching

- Teaching Assistant, *Digital Humans*, ETH Zürich (263-5806-00L) *Spring 2024, 2025, 2026*
- Teaching Assistant, *Computer Vision*, ETH Zürich (263-5902-00L) *Fall 2023, Fall 2024*

Academic Service

- **Co-organizer:** EgoMotion Workshop at ICCV 2025.
- **Reviewer:** CVPR, ICCV, ECCV.