Xinyao Li

[indigolxy | 🌐 mysite | ✉ xinyaoli511@gmail.com

EDUCATION

Shanghai Jiao Tong University(SJTU), Shanghai, China Sep 2022 - Jul 2026

BS in Computer Science, ACM Honor Class (a selective program for 30 students only)

GPA(overall): 89.3/100

Selected Courses: Linear Algebra: 97/100, Principle and Practice of Computer Algorithms: 97/100, Computer Architecture: 95/100, Advanced Compiler Designing: 95/100, Machine Learning: 93/100, Computer Vision: 100/100, Reinforcement Learning: 93/100, Image Generation: 93/100

RESEARCH EXPERIENCE

Anywhere to Sit: Humanoid Whole-body Motion in Restricted Environments

Advisor: Prof. Saurabh Gupta, RoboVision @ Illinois Lab, UIUC, USA Apr 2025 – Present

– Developed a visual-motor tracking policy for agile, contact-rich whole-body motion in constrained spaces, using situations such as entering vehicles and sitting on chairs behind tables.
– Designed scalable reference-trajectory generation and multi-stage RL training pipelines for robust and deployable policy learning.
– Integrated visual inputs via both modular and end-to-end designs to improve generalization across diverse indoor layouts.
– Took a leading role as the primary contributor.

Building a cost-effective and efficient generalizable VLA for robots

Supervisor: Minghuan Liu, Robotics Group, ByteDance Research, Beijing Jan 2025 - Apr 2025

– Achieved comparable performance with a $16\times$ smaller model by distilling pretrained VLA backbone, enabling shorter inference time and lower cost for deployment.
– Distilled pretrained VLA backbone by matching the hidden state latent space in Transformer layers, then finetuned with the action head.
– Validated with the CALVIN benchmark, OXE dataset (for pretraining) and the SIMPLER benchmark.
– Implemented the training and validation infra as the independent developer.

Learning Real-Time Humanoid-Human-Object Interaction from Human Demonstrations

Advisor: Prof. Weinan Zhang, APEX Lab, SJTU, China Jun 2024 - Jan 2025

– Proposed the first real-time humanoid interaction framework capable of learning from human demonstrations, enabling dynamic task-switching and immediate responses to human instructions.
– Trained models directly from human demonstration videos via Human Pose Estimation and Retargeting.
– Developed a fully open-sourced 3D-printed motion capture system for easy reproduction of the pipeline.
– Developed the complete training pipeline and real-world deployment as a primary contributor.
– 2nd author with equal contribution.

SKILLS

Robotics	Simulators(IsaacLab, Mujoco, IsaacGym, etc.), 3D-printing, Real-world Deployment
Control	MPC, robot kinematics, adaptive control
Coding	Python, Pytorch, ROS2, C++