

# LONGFEI LI

loongfei.li03@gmail.com · loongfeili.github.io · Google Scholar · GitHub

## EDUCATION

**Ph.D. Studies**, Beijing Jiaotong University Sep. 2025 – May 2026  
*Information and Communication Engineering* Advisor: Prof. Yunchao Wei

**B.Eng.**, Beijing Jiaotong University Sep. 2021 – Jun. 2025  
*Computer Science and Technology* Rank: 3 / 78

## RESEARCH INTERESTS

My research explores **spatial intelligence** for multimodal and embodied agents, focusing on how models understand 3D structure and spatial relations, transfer these abilities from benchmarks to action, and ground them in **world models and vision-language-action models** for physical environments.

## PUBLICATIONS

**SpatialTree: How Spatial Abilities Branch Out in MLLMs** CVPR 2026 **Highlight**  
*ICLR 2026 Workshop on Efficient Spatial Reasoning* **Best Paper Award** [Project](#)

Yuxi Xiao\*, **Longfei Li**\*, Shen Yan, Xinhang Liu, Sida Peng, Yunchao Wei, Xiaowei Zhou, Bingyi Kang<sup>†</sup>.

Proposed a hierarchical taxonomy to systematically evaluate spatial intelligence in MLLMs from Perception to Action; led benchmark construction and post-training experiments, uncovering asymmetric transfer dynamics between perception and reasoning.

**StereoWorld: Geometry-Aware Monocular-to-Stereo Video Generation** CVPR 2026  
Ke Xing, **Longfei Li**, Yuyang Yin, Hanwen Liang, Guixun Luo, Chen Fang, Jue Wang, Konstantinos N. Plataniotis [Project](#)

Xiaojie Jin, Yao Zhao, Yunchao Wei<sup>†</sup>.

Designed and evaluated components of a geometry-aware monocular-to-stereo generation pipeline for XR/VR content creation.

**Martian World Model: Controllable Video Synthesis with Physically Accurate 3D Reconstructions** NeurIPS 2025 [Project](#)

**Longfei Li**, Zhiwen Fan, Wenyan Cong, Xinhang Liu, Yuyang Yin, Matt Foutter, Panwang Pan, Chenyu You, Yue Wang, Zhangyang Wang, Yao Zhao, Marco Pavone, Yunchao Wei<sup>†</sup>.

Developed a physically grounded world model that reconstructs accurate 3D geometries from Mars rover data and enables controllable video synthesis for robotic simulation.

**AdGPT: Explore Meaningful Advertising with ChatGPT** ACM TOMM 2025 [Project](#)  
Jiannan Huang, Mengxue Qu, **Longfei Li**, Yunchao Wei.

## RESEARCH EXPERIENCE

**Research Intern, ByteDance Seed** Aug. 2025 – Present

- **Aug. 2025 – Mar. 2026**, advised by Dr. Bingyi Kang: Spatial intelligence for MLLMs; co-led *SpatialTree*.
- **Mar. 2026 – Present**, Vision-Language-Action Team: Exploring how multimodal agents can transfer spatial reasoning from benchmarks to action in the physical world through VLA models.

**Research Intern, University of Texas at Austin, VITA Group** Sep. 2024 – Aug. 2025  
*Mentors: Zhiwen Fan, Wenyan Cong*

- Built *Martian World Model*, a geometry-grounded controllable Mars video generator (NeurIPS 2025).

**Research Assistant, Beijing Jiaotong University, WEI Lab** Sep. 2023 – Sep. 2025  
*Advisor: Prof. Yunchao Wei*

- Worked on multimodal understanding, spatial reasoning, and geometry-aware video generation.
- Contributed to *AdGPT* (ACM TOMM 2025) and *StereoWorld* (CVPR 2026).

## HONORS AND AWARDS

Outstanding Graduate of Beijing Jun. 2025  
National Scholarship Oct. 2024  
National Scholarship Oct. 2022