

Personal Statement - Jiahe Xu

A long-term goal drives my work in robotics: to build robots that can learn from humans naturally and operate reliably across different tasks, environments, and hardware. Throughout my experiences in research and industry, I have focused on creating scalable learning pipelines and robust robotic systems that bring embodied AI closer to this vision.

At the CMU Air Lab, I engineered real-time multi-sensor systems, optimized embedded perception pipelines, and developed calibration frameworks that supported the *Flying Hand* and *Flying Calligrapher* projects. These experiences taught me how tightly integrated perception, synchronization, and reliability must be in real-world robotic systems.

In Prof. Katerina Fragkiadaki's lab, I shifted toward large-scale robot learning. I built a complete 3D bimanual pipeline—high-frequency depth estimation, multi-camera streaming, and a rewritten Aloha controller—and benchmarked leading VLA models. This work led to *3D FlowMatch Actor (3DFA)*, a unified 3D manipulation policy that significantly improved accuracy and ran more than 30× faster than prior approaches.

As CTO of PinocchioAI, I applied these ideas to real deployment. I led the development of our navigation robot, built multimodal real-to-sim pipelines, and designed object-centric learning systems for embodiment transfer. Deploying robots in public spaces revealed key weaknesses of current VLA methods—poor robustness, limited transferability, and costly data collection—and motivated my work on OC3D, a framework addressing these issues through object-centric 3D diffusion.

Across these projects, I've developed a clear research direction: scalable robot learning requires standardized pipelines that unify perception, 3D representation, real-to-sim data, and human demonstrations. Through a Ph.D. program, I hope to build policies that generalize across hardware generations, learn efficiently from people, and enable robots to perform reliably in diverse settings.

I am motivated to continue pursuing this vision and contribute to the next generation of embodied AI systems.