Yiteng Xu

Computer Vision and Robotics

(+86) 13130463002
xuyt2023@shanghaitech.edu.cn
yiteng-xu.github.io/
Google Scholar

Education

Sep. 2023 – Present **5-Year Master-PhD Program (3rd Year)**, *ShanghaiTech University*, Shanghai
Computer Science and Technology

Sep. 2019 – Jun. 2023 **Bachelor of Science**, *ShanghaiTech University*, Shanghai
Computer Science and Technology

Core Strengths and Skills

Research Impact First/co-first author of papers published at top-tier CV/AI conferences including **CVPR**, **AAAI (Oral)**, and **ICCV**, with proven ability to define novel problems and develop SOTA solutions.

Core Expertise
- **3D Perception and Reconstruction (Expert):** Proficient in **LiDAR & Camera** sensor fusion for real-time 3D scene and human reconstruction using sparse/dense data.
- **Embodied Intelligence and Interaction (Expert):** Focus on humanoid robots with expertise in **multi-agent interaction**, **human-robot collaboration**, and complex behavior generation.
- **Parametric Human Modeling (Expert):** Expert in **SMPL/SMPL-X** from model fitting and motion generation to integration with physics-based simulation.
- **3D Motion Generation and Prediction (Advanced):** Familiar with generative models (Diffusion, Auto-regression) to synthesize realistic, physically-plausible human motion.

Technical Stack
- **Programming and Algorithms (Expert): Python**, **PyTorch**, with solid knowledge of algorithms and data structures.
- **3D Vision and Simulation (Advanced): NVIDIA Isaac Gym/Orbit**, Open3D, PCL, PyTorch3D, Trimesh, MeshLab for geometry processing, physics simulation, and visualization.
- **Data Science (Advanced):** Pandas, NumPy, SciPy, Scikit-learn, Matplotlib/Seaborn for multimodal temporal data processing and analysis.
- **DevOps (Expert): Linux**, **Git/GitHub**, **Slurm**, **Docker**, **Bash**; skilled in managing scalable environments and cluster-based automation.

Methodology and Soft Skills
- **System Architecture (Advanced):** Designed a closed-loop simulation platform integrating AR, motion capture, and physics-based learning from scratch.
- **User-Centered Experiment Design (Advanced):** Designed "Wizard of Oz" and Human-in-the-Loop studies to collect high-quality interactive datasets.
- **Technical Communication and Leadership (Expert):** Proficient in LaTeX, Draw.io; skilled in writing, project planning, and team collaboration.

Research Experience

SymbioSim: A Human-in-the-Loop Simulation Platform for Bidirectional Learning, *Co-first Author, Under Review:* **SIGGRAPH 2025**
- Proposed **SymbioSim**, a novel simulation platform enabling **continuous, bidirectional learning** between humans and robots, addressing the lack of authentic feedback in conventional simulators.
- Developed a full-stack architecture combining **AR interaction**, **real-time LiDAR-based motion capture**, and physics-based simulation (Isaac Gym), achieving automated model refinement via real-world interaction.
- Validated through user studies that the platform not only allows **robot learning from feedback**, but also helps users **gradually adapt to and trust robots**, advancing symbiotic intelligence.

UniPVU-Human: A Unified Framework for Human-Centric Point Cloud Video Understanding, *First Author,* **CVPR 2024**

- Proposed a unified and efficient point-cloud video understanding framework for dynamic human understanding in robotics and autonomous driving.
- Designed hierarchical self-supervised learning to reduce labeling cost by **70%**, achieving high-quality human semantics and dynamics without manual annotations.
- Introduced semantic-guided architecture and lightweight distillation, reducing model memory by **65%** while achieving SOTA performance (+3.8% in action recognition).

Weakly-Supervised Multimodal 3D Human Pose Estimation, *Co-first Author,* **AAAI 2023 (Oral)**, **ICCV 2023**

- Proposed IPAFusion, a cross-modal attention fusion mechanism aligning image and LiDAR features **without precise calibration**, improving robustness by **37%**.
- Developed self-supervised framework leveraging geometry and temporal constraints, **reducing 3D annotation cost by over 90%** and validated in **70m-scale** outdoor settings.
- Constructed HuCenLife, the first multimodal dataset for multi-agent human-robot interaction, serving as a benchmark for Embodied AI and autonomous driving.

Honors and Awards

National Scholarship (Top 1%)

Shanghai Outstanding Graduate

2nd Prize, National Robotics Competition

3rd Prize, Innovation and Entrepreneurship Competition, ShanghaiTech University

Merit Student & Outstanding Student Leader, ShanghaiTech University