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**, 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).

Human-centric Scene Understanding for 3D Large-scale Scenarios, *Co-first Author,* ***ICCV 2023***

- Pioneered and constructed **HuCenLife**, the first large-scale, multimodal dataset focused on complex human-centric interactions, addressing a critical data gap for Embodied AI.
- Led the **end-to-end data pipeline**: from designing the capture system (**128-beam LiDAR, 6 cameras**) to managing the fine-grained annotation of over **6,000 frames** across 32 diverse scenes.
- Processed and structured a massive dataset containing **65k+ human instances** and **31k+ interacted objects**, establishing a foundational benchmark for robust 3D perception research (e.g., segmentation, detection, action recognition).

Weakly-Supervised 3D Human Pose Estimation in Large-Scale Scenes, *Co-first Author,* ***AAAI 2023 (Oral)***

- Invented **IPAFusion**, a novel cross-modal attention mechanism that automatically aligns image and LiDAR features **without requiring precise sensor calibration**, overcoming a major hurdle for real-world deployment.
- Developed a weakly-supervised learning framework that leverages geometric and temporal constraints, **eliminating the need for expensive 3D annotations** and reducing labeling costs by over **90%**.
- Validated the system's scalability and accuracy in complex, **large-scale outdoor environments (up to 70m range)**, proving its practical value for industrial robotics and autonomous systems.

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

Skill Points

3D Computer Vision, Robotics, Embodied AI, Human-Robot Interaction (HRI), 3D Perception, Human-centric Scene Understanding, Parametric Human Modeling (SMPL/SMPL-X), 3D Motion Generation and Prediction, Humanoid Robotics, Multi-Agent Interaction, Physics-based Simulation and Modeling, Sensor Fusion (LiDAR & Camera), Human-in-the-Loop (HITL) Simulation, Augmented Reality (AR) in Robotics, Real-time Motion Capture, Weakly Supervised Learning, Self-Supervised Learning for Robotics, Geometric Deep Learning, 3D Human Pose Estimation, Point Cloud Video Understanding, Assistive Robotics, Python, PyTorch, NumPy, Pandas, SciPy, Scikit-learn, Matplotlib, Open3D, PCL, PyTorch3D, Trimesh, MeshLab, LaTeX, Bash Scripting, Algorithms and Data Structures, Docker, Slurm, Git / GitHub, PyTorch Lightning, OpenCV, Jupyter Notebook, Linux, Deep Learning, Generative Models, Diffusion Models, Autoregressive Models, Transformers, 3D Point Cloud Processing, LiDAR Data Processing, Point Cloud Segmentation (Semantic/Instance), Point Cloud Denoising, 3D Action Recognition, 3D Scene Flow Estimation, 3D Human Mesh Recovery, 3D Data Annotation, Inverse Kinematics (IK), Forward Kinematics, Geometric Constraints Modeling, Spatio-temporal Representation Learning, Chamfer Distance, Point-based models (PointNet, PointNet++), 3D Visualization, Robot Perception Systems, Robot Learning, Motion Planning and Control, Behavior Generation, Human-Robot Collaboration, Real-time Interactive Systems, Robot Simulation Environments, Sim-to-Real Transfer, Augmented Reality Interfaces for Robotics, Multi-robot Systems, Data Science, Multimodal Temporal Data Processing, Data Analysis, Data Visualization, Large-scale Dataset Management, High-Quality Interactive Dataset Collection, Closed-loop System Integration, Real-time Data Streaming, Scientific Writing, Project Planning and Management, Problem Definition and Formulation, User-Centered Experiment Design, Human-in-the-Loop Studies, "Wizard of Oz" Prototyping, Quantitative and Qualitative Research, Technical Documentation (LaTeX, Draw.io).