YitengXu

Embodied AI and Human-Robot Symbiosis

(+86) 13130463002
xuyt2023@shanghaitech.edu.cn
yiteng-xu.github.io/
yiteng-xu-9433a4371/
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 HRI/CV/AI conferences including **SIGGRAPH Asia**, **CVPR**, **AAAI (Oral)**, and **ICCV**, with proven ability to define novel problems and develop SOTA solutions.

Core Expertise

- **Human-Robot Symbiosis (Expert):** Focused on the **bidirectional, adaptive learning** between humans and humanoid robots. Expert in designing closed-loop **Human-in-the-Loop (HITL)** systems to build self-evolving, "better-with-use" intelligent agents.
- **Cyber-Physical Interaction (Expert):** In-depth experience in fusing **Augmented Reality (AR)** with the physical world. Proficient in leveraging **real-time LiDAR motion capture** to create natural and immersive HRI experiences.
- **3D Perception & Parametric Modeling (Expert):** Proficient in **LiDAR & Camera** fusion for reconstruction. Expert in **SMPL/SMPL-X** covering the full pipeline from motion capture and generation to integration with physics-based simulation.
- **3D Interactive Motion Generation (Advanced):** Familiar with generative models (e.g., Diffusion) to synthesize high-quality, physically plausible, and socially-aware interactive human/robot motions.

Technical Stack

- **Programming and Algorithms (Expert): Python**, **PyTorch**, with solid knowledge of algorithms and data structures.
- **3D Vision & Simulation (Advanced): NVIDIA Isaac Gym/Orbit**, Open3D, PCL, PyTorch3D; Hands-on experience with hardware integration including **PICO AR glasses** and **Ouster LiDAR**.
- **Data Science (Advanced):** Pandas, NumPy, SciPy, Scikit-learn, Matplotlib/Seaborn for multimodal temporal data processing and user behavior analysis.
- **DevOps (Expert): Linux**, **Git/GitHub**, **Slurm**, **Docker**, **Bash**; skilled in managing scalable environments and cluster-based automation.

Methodology and Soft Skills

- **Complex System Architecture (Expert): Architected and implemented SymBridge**, a complex closed-loop system integrating AR, real-time LiDAR MoCap, physics simulation, and real-world robotics, demonstrating exceptional system-level design and execution capabilities.
- **User-Centered Experiment Design (Advanced):** Successfully designed and executed a **50-participant** Human-in-the-Loop user study, skilled in collecting and analyzing high-quality interaction data through scientific methodologies.
- **Technical Communication and Leadership (Expert):** Proficient in LaTeX, Draw.io; skilled in scientific writing, project planning, and team collaboration.

Research Experience

SymBridge: A Human-in-the-Loop Cyber-Physical System for Adaptive Human-Robot Symbiosis, *Co-first Author (Lead on System Architecture & Core Algorithm),* ***SIGGRAPH Asia 2025 (Oral)***

- **Pioneered SymBridge**, a novel Cyber-Physical system that solves the dual challenges of lacking authentic human data in simulation and the high cost of real-robot testing, enabling safe, efficient, and **bidirectional adaptive learning**.
- **Architected the end-to-end system**, integrating **AR glasses (PICO 4)** for immersive interaction, **real-time LiDAR (Ouster)** for human perception, and **physics simulation (Isaac Gym)** for policy generation. Successfully **deployed the learned model to a real humanoid robot (LEJU Kuavo)**.
- **Proposed a novel robotic interaction model** and validated it via a large-scale **50-participant user study**, proving the system enables robots to **learn from human feedback** while helping humans **adapt to and trust the robot**, paving the way for true human-robot symbiosis.

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.

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.

Patents

2025.04 **A General Model for Dynamic Point Cloud Understanding and Multi-Task Collaborative Optimization**, *Inventors:* ***Yiteng Xu****, Yuesuo Ma*, Publication No.: CN120412083A

- Proposed a unified framework to efficiently and accurately interpret human dynamics in complex scenes through multi-level feature fusion and temporal model optimization.
- The technology supports applications in **intelligent surveillance, VR/AR, human-computer interaction, autonomous driving**, and **sports analytics**.

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 Understand-

ing, 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).