

## 具身智能与人机共生系统方向

## 教育背景 (Education)

- Sep. 2023 – Present 五年制硕博连读 (在读三年级), 上海科技大学, 上海  
计算机科学与技术
- Sep. 2019 – Jun. 2023 本科, 上海科技大学, 上海  
计算机科学与技术

## 核心优势与专业技能 (Skills)

**研究影响力** 作为一作/共同一作在 HRI/CV/AI 领域顶会 SIGGRAPH Asia, CVPR, AACL (Oral), ICCV 发表多篇论文, 具备定义前沿问题与实现 SOTA 解决方案的成熟能力。

## 核心专业领域

- **人机共生系统 (精通)**: 聚焦人形机器人与人的双向、自适应学习。精通设计闭环“人在环路”(HITL) 系统, 以构建自演化、“越用越好”的智能体。
- **虚实融合交互 (精通)**: 在融合增强现实 (AR) 与物理世界方面具备深厚经验。熟练运用实时 LiDAR 动捕技术, 以创造自然且沉浸式的人机交互体验。
- **3D 感知与参数化建模 (精通)**: 熟练掌握 LiDAR 与相机的融合重建技术。精通 SMPL/SMPL-X 模型, 覆盖从动捕、生成到与物理仿真集成的全链路应用。
- **3D 交互行为生成 (熟练)**: 熟悉使用生成模型 (如 Diffusion) 合成高质量、符合物理规律且具备社交智能的交互式人体/机器人运动。

## 技术栈与工程能力

- **编程与算法 (精通)**: Python, PyTorch; 具备扎实的算法与数据结构知识。
- **3D 视觉与仿真 (熟练)**: NVIDIA Isaac Gym/Orbit, Open3D, PCL, PyTorch3D; 拥有包括 PICO AR 眼镜和 Ouster LiDAR 在内的软硬件集成实战经验。
- **数据科学 (熟练)**: Pandas, NumPy, SciPy, Scikit-learn, Matplotlib/Seaborn, 用于多模态时序数据处理和用户行为分析。
- **开发与运维 (精通)**: Linux, Git/GitHub, Slurm, Docker, Bash; 擅长管理可扩展开发环境与集群自动化。

## 研究方法 with 软实力

- **复杂系统架构 (精通)**: 主导架构设计并实现了 SymBridge——一个集成了 AR、实时 LiDAR 动捕、物理仿真和真实世界机器人的复杂闭环系统, 展现了卓越的系统级设计与执行能力。
- **人本位实验设计 (熟练)**: 成功设计并执行了一项 50 名参与者的“人在环路”用户研究, 擅长通过科学方法论收集并分析高质量的交互数据。
- **技术沟通与领导力 (精通)**: 熟练运用 LaTeX, Draw.io; 擅长科学写作、项目规划和团队协作。

## 研究与项目经历 (Research Experience)

**SymBridge**: 面向自适应人机共生的“人在环路”虚实交互系统, 共同一作 (系统架构与算法核心贡献者), SIGGRAPH Asia 2025 (Oral)

- **首创 SymBridge**, 一个新颖的虚实结合系统, 解决了仿真中缺乏真实人类数据和真实机器人测试成本高昂的双重挑战, 实现了安全、高效的双向自适应学习。
- **主导设计了端到端系统架构**, 集成了 AR 眼镜 (PICO 4) 用于沉浸式交互、实时 LiDAR (Ouster) 用于人体感知以及物理仿真 (Isaac Gym) 用于策略生成。成功将学习到的模型部署到真实人形机器人 (LEJU Kuavo)。
- **提出了一种新颖的机器人交互模型**, 并通过一项大规模的 50 人用户研究进行验证, 证明该系统能使机器人从人类反馈中学习, 同时帮助人类适应并信任机器人, 为实现真正的人机共生铺平了道路。

### UniPVU-Human: 面向人体的通用点云视频理解统一框架, 独立一作, *CVPR 2024*

- 针对机器人与自动驾驶中的动态人体理解, 提出了一个统一且高效的点云视频理解框架。
- 设计了分层自监督学习方法, 将标注成本降低 70%, 在无需人工标注的情况下实现了高质量的人体语义与动态特征提取。
- 引入了语义引导架构和轻量级蒸馏技术, 将模型内存占用降低 65%, 同时在动作识别等任务上取得了 SOTA 性能 (提升 3.8%)。

### 面向 3D 大规模场景的以人为中心的场景理解, 共同一作, *ICCV 2023*

- 开创并构建了 HuCenLife, 首个专注于复杂人本交互的大规模、多模态数据集, 解决了具身智能领域的关键数据缺口。
- 主导了端到端数据管线: 从设计采集系统 (128 线 LiDAR, 6 个摄像头) 到管理 32 个不同场景下超过 6,000 帧的精细化标注。
- 处理并构建了一个包含超过 6.5 万个人体实例和超过 3.1 万个交互物体的海量数据集, 为鲁棒的 3D 感知研究树立了基础性基准。

### 弱监督大规模场景下 3D 人体姿态估计, 共同一作, *AAAI 2023 (Oral)*

- 发明了 IPAFusion, 一种新颖的跨模态注意力机制, 无需精确传感器标定即可自动对齐图像和 LiDAR 特征, 克服了现实世界部署的一大障碍。
- 开发了一个弱监督学习框架, 利用几何与时间约束, 消除了对昂贵 3D 标注的需求, 将标注成本降低了 90% 以上。
- 在复杂的大规模室外环境 (范围达 70 米) 中验证了系统的可扩展性与准确性, 证明了其在工业机器人和自动驾驶系统中的实用价值。

---

## 发明专利 (Patents)

2025.04    动态点云通用理解模型及多任务协同优化系统, 发明人: 许艺腾, 马月所, 公开号: CN120412083A

- 提出了一种统一的动态点云理解框架, 通过多层次特征融合与时序模型优化, 旨在高效、精确地理解复杂场景中的人体动态。
- 该发明可广泛应用于智能监控、虚拟现实 (VR/AR)、人机交互、自动驾驶及体育分析等前沿领域, 为相关技术提供了核心算法支持。

---

## 荣誉奖项 (Honors and Awards)

国家奖学金 (Top 1%)

上海市优秀毕业生

第十九届全国大学生机器人大赛, 全国二等奖

上海科技大学第六届创新创业大赛, 三等奖

上海科技大学三好学生、优秀学生干部