Xiangyu CHEN

(+86) 133-9340-2613 | changerc77@gmail.com | changerc77.github.io | ChangerC77

EDUCATION

The Hong Kong University of Science and Technology (Guangzhou) *Guangzhou, China*
MPhil in ROAS Thrust, co-advised by **Prof. Renjing Xu** and **Prof. Haoang Li** *Sep. 2025 – Jun. 2027*

Liverpool John Moores University *Liverpool, UK*
BEng Electrical and Electronic Engineering (First Class Honours) *Sep. 2019 – Jun. 2023*

RESEARCH INTERESTS

Embodied AI, Manipulation, Data Collection, VLA, Tactile, Mapping and Navigation

PUBLICATIONS

- **Beyond Viewpoint Generalization: What Multi-View Demonstrations Offer and How to Synthesize Them for Robot Manipulation?**
  Boyang Cai*, Qiwei Liang*, Jiawei Li*, Shihang Weng*, Zhaoxin Zhang*, Tao Lin, **Xiangyu Chen**, Wenjie Zhang, Jiaqi Mao, Weisheng Xu, Bin Yang, Jiaming Liang, Junhao Cai, Renjing Xu
  Under Review
- **RoboGEM: Learning Language-guided Robotic Manipulation via Generalizable and Efficient Feature Distillation**
  Chunzheng Wang, Yuhang Zheng, **Xiangyu Chen**, Weize Li, Songen Gu, Yupeng Zheng
  ACM Multimedia, RoboSoft'25 Workshop (Oral) Best Paper Finalist, 2025
- **GaussianGrasper: 3D Language Gaussian Splatting for Open-vocabulary Robotic Grasping**
  Yuhang Zheng, **Xiangyu Chen**, Yupeng Zheng, Songen Gu, Runyi Yang, Bu Jin, Pengfei Li, Chengliang Zhong, Zengmao Wang, Lina Liu, Chao Yang, Dawei Wang, Zhen Chen, Xiaoxiao Long, Meiqing Wang
  RAL, 2024
- **Block-Map-Based Localization in Large-Scale Environment**
  Yixiao Feng, Zhou Jiang, Yongliang Shi, Yunlong Feng, **Xiangyu Chen**, Hao Zhao, Guyue Zhou
  ICRA, 2024

RESEARCH EXPERIENCE

BrainCo Technology Co., Ltd. *Shenzhen, China*
RESEARCH INTERN | ADVISOR: JIA ZUO *Dec. 2025 – Present*

- **Dexterous Hand Exoskeleton Data Collection Development**

  Introduction: We propose a *DexUMI-based multi-modality data collection system* to bridge the *proprioceptive and cross-embodiment gaps* in dexterous manipulation. By shifting from traditional UMI interfaces to an *integrated exoskeleton system*, we streamline the collection of complex human-like demonstrations while ensuring *high-fidelity sensory-motor alignment*.
  Dexterous Hand Design: Developed a *DexUMI-based* multimodal data acquisition framework specifically tailored for *21-DoF dexterous hands*. This system implements a high-efficiency pipeline to synchronously capture *robot actions, dexterous hand states, RGB video streams, and tactile feedback*. By utilizing a *human-centric retargeting algorithm*, the platform translates intricate finger kinematics into robot action spaces in *real time*, enabling efficient data collection for complex manipulation tasks.
  6-DoF Pose Tracking: Conducted a comparative analysis of localization modalities (*Mocap, Infrared, and VIO*) to identify the optimal solution for *in-the-wild* applications. By benchmarking these methods against environmental constraints, we implemented a *self-contained VIO-based pipeline* to ensure robust, infrastructure-free tracking *in the wild*, effectively overcoming the limitations of external reference systems in unstructured settings.

Humanoid Computing Lab(HCLab) (collaborate with NVIDIA) *Guangdong, China*
ADVISOR: PROF. RENJING XU *Dec. 2025 – Present*

- **Sim2Real Multi-Modality Data Collection System**

  Introduction: To bridge the *sim-to-real* gap in *visuo-tactile multimodal* models and address the dearth of *standardized benchmarks*, we propose an efficient *UMI-based* data acquisition framework alongside an *end-to-end sim-to-real pipeline* for data collection and model training. This approach aims to facilitate *seamless policy transfer* and provide a robust evaluation environment for complex manipulation tasks

UMI-based Design: Led the development of a *low-cost UMI gripper platform* that dramatically lowers the financial barrier for data collection. It features *real-time synchronization* between the UMI and the robot, integrating *multimodal tactile and force data* to mimic human-like manipulation. The inclusion of a precision *force-controlled data collection method* further optimizes efficiency, ensuring *high-quality datasets* for model training.
Sim2Real Transfer: Developed a *multi-modality data alignment pipeline* to bridge real-world data with *Isaac Lab simulation platform*. This workflow facilitates the transformation and alignment of *raw multimodal data* into the *LeRobot format*, establishing a seamless mapping between *physical sensor telemetry* and simulated data for consistent policy evaluation

TARS Technology (Shanghai) Co., LTD. — *Shanghai, China*

Research Intern | Advisor: **Prof. Yilun Chen** and **Prof. Wenchao Ding** — *Mar. 2025 – Dec. 2025*

- **SenseHub Human-Centric Multi-Modal Data Collection System Development**

  Introduction: SenseHub is a *human-centric multimodal data collection system* that provides the empirical foundation for the *WIYH* dataset and the *AWE* foundation model through *real-world data acquisition*.
  Teleoperation System: Engineered a comprehensive *teleoperation and data acquisition system* from the ground up, supporting *SpaceMouse, VR, and isomorphic arm* control to facilitate *large-scale, multimodal robotic data collection*.
  UMI-based Data Collection System: *Led* the development of the initial *UMI-based* hardware data collection framework, integrating *dexterous hands* and *grippers* with *multi-modal sensors* to capture high-quality *visual, tactile, and force feedback* for *Embodied AI* training.
  Tactile Sensors Survey: Conducted *comprehensive benchmarking and testing* across multiple tactile sensing modalities—including *vision-based, piezoresistive, capacitive, magnetic, and Hall-effect sensors*—to select the optimal hardware solution tailored for specific research objectives.
  Results: *Led* the development of *teleoperation, UMI-based data collection infra, and real-world robot infra*, providing the *foundational experimental platform* for several high-impact research papers, including *OmniVTA* and *World In Your Hands (WIYH)*

- **The First Prize in Google DeepMind & ICRA 2025 WBCD Challenge (Team Leader)**

  Introduction: This challenge is jointly organized by Google DeepMind and ICRA, aiming to propose efficient data collection strategies and exploring the benchmark of *bimanual manipulation* in real-world scenarios, with evaluation metrics including *execution speed, system reliability, and policy implementation*.
  Bimanual Manipulation System Implementation: Design a *mobile manipulation system* to efficiently collect real data and robustly perform *tabletop manipulation and mobile tasks*. *Dual-arm* can do tasks including unfold tablecloth, open the lid, place the pizza and close the lid tasks while mobile. Designed grippers can grasp cups and open boxes, etc.
  Data Collection: Benchmark data collection methods, such as *teleoperation* (e.g., VR and hand-held grippers) and *learning from videos* methods, propose methods suitable for collecting data in real scenarios.
  Design Hierarchical Bimanual Manipulation Policy for Dining Room Service: Propose *Bimanual Manipulation Models (VLA)* to execute *long-horizon*, complex and fine mobile manipulation tasks in real-world dining room service scenarios, with focus on designing models that can mitigate *compounding errors* and improve robustness against *dynamic interference*.

- **Champion in ManiSkill-ViTac 2025 Track 2**

  Introduction: The competition evaluated robots manipulation capabilities in complex environments, requiring them to determine object shapes and select correct insertion slots by integrating *tactile and visual sensory inputs*.
  Methods: The solution combined *tactile encoders* with *3D point cloud encoding*, enhanced by *positional embeddings*, and employed a *conditional diffusion policy* for robust multi-modal action planning.
  Contribution: Implement *sim-to-real transfer* in real-world experiments, process *tactile sensor data*.

EncoSmart Technology (Beijing) Co., LTD. — *Beijing, China*

Research Intern | Advisors: **Prof. Xiaoxiao Long** — *Nov. 2023 – Feb. 2024, Jul. 2024 – Dec. 2024*

- **GaussianGrasper: 3D Language Gaussian Splatting for Robotic Grasping**

  3D Reconstruction based Manipulation Design: proposed a method utilizes *3D Gaussian Splatting* to explicitly represent the scene as a collection of *Gaussian primitives*. The method takes fewer views *RGB-D* to solve the inconsistency issue between *geometry and semantic information* in vision-based manipulation, as well as the inability to dynamically update the scenes.
  Data Collection: *Hand-eye calibration* of robotic arm and camera is performed to improve the accuracy of *multi-view point cloud fusion* during data collection
  Method Design: Using *visual language models (VLMs)* to achieve scene understanding and *object grounding*, and employ Grasp Models (*AnyGrasp*) to generate the *6-DoF pose* for grasping.
  Experiment: Design strategies to operate robotic arm and gripper, enabling *long-horizon and continuous manipulation*, as well as update scenes.

GAIRLAB, City University of Hong Kong — *Hong Kong, China*

Research Intern | Advisors: **Prof. Peng Yin** — *Mar. 2024 – Jun. 2024*

- **Mobile Manipulation Policy based on Imitation Learning**

  Learning-based Mobile Manipulation System Design: Proposed robot policies based on the *ACT network* for *mobile manipulation* to improve the robustness in performing *long-horizon and fine manipulation tasks*.
  Data Collection: Data collection by using *whole-body teleoperation system*.

Method Design: Increase the degrees of freedom of whole-body control to *17 DoF* and optimize the control strategy through position control and *object-centered coordinate relationships* to address the high task failure rate.
Navigation Module Design: Focus on *learning-based navigation* method and how to leverage *VLMs* to improve the *indoor visual navigation* efficiency and localization accuracy.

AIR, Tsinghua University & ARXRobotics *Beijing, China*
RESEARCH INTERN @ **DISCOVER LAB** | ADVISOR: **PROF. GUYUE ZHOU**, **DR. YONGLIANG SHI** AND **PROF. JIANGTAO GONG** *Sep. 2021 – Nov. 2023*

- **Block-Map-Based Localization in Large-Scale Environment** *Sep. 2021 – Nov. 2023*
  Block Maps Localization System Design: Proposed a *subgraph localization system* based on generating *block maps* and corresponding *switching strategies*. The method can address the issue of easy loss of robot localization, improving *localization accuracy by at least 3 times* while increasing *computational speed by 150%*.
  Perception Module Design: Used *multi-line LiDAR* and *LIO-SAM* algorithms to create maps in large-scale scenes. Used *C++* and *PCL point cloud library* to downsample point clouds for removing ground points.
  Navigation Module Design: Using the *A* algorithm* and *TEB algorithm* as planners to enable robots to achieve *dynamic obstacle avoidance*. These planners assist robots in conducting localization experiments.
- **Multi-Agent Swarm Formation Navigation Algorithms** *May 2022 – Feb. 2023*
  Multi-Agent Navigation Strategies: Proposed a *multi-agent collaboration system* to improve work efficiency through cooperation of multiple robots.
  Method Design: Develop *Leader-follower*, *artificial potential field* and *pure pursuit* algorithms to enable robots to flexibly achieve *single-agent and multi-agent collaborative obstacle avoidance*.
  Sim2Real Pipeline: Use *Gazebo* and *Isaac Sim* simulators to test algorithms in simulation, then algorithms are fine-tuned and deployed to the *real robots* to ensure the desired instantaneity and robustness for multi-agent obstacle avoidance.
- **Obstacle Avoidance Algorithm for Indoor Racing Unmanned Vehicles** *May 2022 – Feb. 2023*
  Motivation and hardware: We aim to develop navigation system that can improve the robot's localization and real-time obstacle avoidance robustness in corridor and glass environments. The hardware includes *single-line Lidar, IMU and depth camera*, with *differential drive model* as the kinematic model.
  SLAM Experiment: We benchmark the performance of *Lidar-based SLAM* (*Gmapping, Cartographer*) algorithms and *vision-based SLAM* algorithms (*Rtabmap, ORB_SLAM*) and design a *multi-sensor fusion algorithm* utilizing *particle filtering, extended Kalman filtering and graph optimization* to improve real-time localization accuracy during motion.
  Navigation Experiment & Results: We benchmarked path planning algorithms based on *graph search and sampling methods*. Experiments revealed that using the *A* algorithm* and *TEB algorithm* as planners outperformed other methods in avoiding *static and dynamic obstacles* during long-horizon navigation, with robots successfully avoiding all tested obstacles
- **End-to-end Visual Navigation based on Reinforcement Learning** *Jun. 2021 – Sep. 2021*
  Motivation: We explored the performance of *end-to-end vision-based navigation* for robots, drawing inspiration from the end-to-end approach used in autonomous driving.
  Pipeline Design: We used a *CNN network* as the backbone for *feature extraction* and collected camera data and control data in a *simulator* to create the dataset. After fine-tuning, the system was deployed on a robot for *real-world experiments*. Results showed that the system successfully completed all obstacle avoidance tasks during the day, but many tasks failed at night due to factors such as low lighting.

HONORS & AWARDS

INTERNATIONAL

2025	**Champion**, ICRA 2025 What Bimanuals Can Do (WBCD) Challenge	*Atlanta, U.S.A*
2025	**Champion**, ManiSkill-ViTac 2025: Challenge on Manipulation Skill Learning With Vision and Tactile Sensing	*Beijing, China*

DOMESTIC

2023	**City Special Prize**, "Unbounded·2023 Shanghai International Student - (Nationalized University Students) Innovation and Entrepreneurship Competition"	*Shanghai, China*
2017	**First Prize (National Level)**, 2017 International Youth Innovation Design Competition China Region	*Beijing, China*

SKILLS

Robotics Technology	ROS1/2
Software Technology	C/C++, Python, MATLAB, OpenCV, Linux and Git, PyTorch
Hardware Technology	Embedded Development (STM32, ESP32), SolidWorks, PCB Design
Language Proficiency	Mandarin (Native), English (Working Proficiency)