LIN DIXUAN

dixuan@seas.upenn.edu Google Scholar github Website

EDUCATION

University of Pennsylvania, USA *Aug. 2024 - May. 2026*
Master of Robotics
GPA: 3.9/4.0

Sun Yat-Sen University, China *Sep. 2019 - Jun. 2023*
Bachelor of Software Engineering
GPA: 3.9/4.0

RESEARCH EXPERIENCE

University of Pennsylvania | GRASP *Nov. 2024 - Nov. 2025*
Master Thesis
- Supervisor: Prof. Kostas Daniilidis
- Motion generation, Scene reconstruction

Tsinghua University | Department of Automation *Jul. 2023 - Jul. 2024*
Research Assistant
- Supervisor: Prof. Yebin Liu
- Temporal 3D Hand Shape Reconstruction, Multi-view 3D Reconstruction

Sun Yat-sen University | ISEE Lab *Sep. 2022 - May 2023*
Undergraduate Research Assistant
- Supervisor: Prof.Wei-shi Zheng
- Person Identification, Multi-Modality Alignment, Image Generation

PROJECTS

Person Re-Identification based on multi-modal information

- The goal of this project is text-based person re-identification in RGB images: retrieving a target individual from natural-language descriptions.
- Distinguishes and jointly learns the two non-equivalent directions: text to image and image to text, instead of collapsing them into a single association.
- Uses text tokens as anchors to adaptively select and aggregate relevant image patches, passing guidance from text to vision.
- Associates image regions with text attributes/phrases by locating supporting regions and reconstructing masked phrases via masked attribute modeling.
- Tests were carried out on three mainstream benchmark datasets CUHK-PEDE, ICFG-PEDE, and RSTPReid. The approach significantly outperformed existing State-of-the-Art (SOTA) across all evaluation metrics.

Reconstructing Interactive Hands in 4D with Transformers

- We aim to build a robust 4D hand-mesh reconstruction system that accepts single images, videos, or multi-view inputs and produces high-quality hand meshes.
- We introduce a Relation-aware Two-Hand Tokenization (RAT) in our transformer. By embedding positional relation information in the tokenization process, RAT enables a generic and informative process of hand inputs across diverse scenarios, regardless of whether the two hands are interacting or separated.
- We propose a 4D Interaction Reasoning (FIR) module to fuse spatial or temporal hand features. The collaboration of RAT and FIR further enhances the stability of hand mesh and temporal movement recovery.

- Our final solution, OmniHands, achieves robust hand mesh recovery in real-world scenarios, serving as a versatile solution to handle hand inputs in various forms, whether the images contain single or two hands, in single frames or temporal sequences or multi-view sequences.

Zero-shot hand-object interaction generation.

- The objective of our project is to synthesize complete scenes with realistic hand–object interactions, conditioned on a single object-centric image and a textual prompt.
- We leverage third-party repositories SAM, InstantMesh, MASTR3, and DepthPro to reconstruct and align scene and object meshes. We then use KlingAI to synthesize plausible hand–object interactions for motion conditioning.
- We train a motion prior on the GRAB dataset that generates high-quality hand–object motions conditioned on 2D keypoints.

PUBLICATIONS

1. **OmniHands: Towards Robust 4D Hand Mesh Recovery via A Versatile Transformer**
   Arxiv. preprint (2025) [paper] [project]
   Dixuan Lin, Yuxiang Zhang, Mengcheng Li, Yebin Liu, Wei Jing, Qi Yan, Qianying Wang*, Hongwen Zhang*
2. **Cross-Modal Adaptive Dual Association for Text-to-Image Person Retrieval**
   IEEE Transactions on Multimedia (2024) [paper]
   Dixuan Lin, Yixing Peng, Jingke Meng*, Wei-Shi Zheng

AWARD

The Chinese Mathematics Competitions
The second prize of province
China Undergraduate Mathematical Contest in Modeling
The first prize of province

SKILLS

- Proficient in Python programming, able to implement ideas through code, and have experience in reproducing models from high-quality conference papers.
- Familiar with commonly used techniques and principles of deep neural networks.
- Skilled in using LaTeX, with experience in academic paper writing, capable of independently searching, reading, and summarizing academic papers.