

Yiming Zhong

Shanghai, China

Email: zhongym@shanghaitech.edu.cn

Mobile: +86 19821234507

Homepage: ymzhong66.github.io

EDUCATION

- **ShanghaiTech University** *Sept. 2024 – Jun. 2027*
M.S. in Computer Science
Advisors: Prof. Yuexin Ma, Prof. Xinge Zhu
- **Shandong University** *Sept. 2020 – Jun. 2024*
B.E. in Statistics; Second Major in Computer Science

RESEARCH INTERESTS

My research focuses on **Embodied Intelligence**, with an emphasis on learning predictive world representations that enable generalizable perception, reasoning, and action in complex physical environments.

- **World and Action Modeling**: Learning action-conditioned predictive representations from multimodal observations and interactions for visuomotor control and long-horizon decision-making.
- **Generalizable Robot Learning**: Developing scalable learning frameworks that transfer across tasks, embodiments, environments, and previously unseen scenarios.
- **Spatial and 3D Intelligence**: Learning geometry-aware representations that integrate visual and generative priors for spatial reasoning, reconstruction, and embodied interaction.

PUBLICATIONS

First-Author / Project-Lead Publications

* Equal contribution; † Corresponding author; ‡ Project lead.

1. **Yiming Zhong***, Qi Jiang*, Jingyi Yu, Yuexin Ma†. **DexGraspAnything: Towards Universal Robotic Dexterous Grasping with Physics Awareness.** (CVPR 2025 [Highlight])
2. Yufei Zhu*, **Yiming Zhong***, Zemin Yang, Peishan Cong, Jingyi Yu, Xinge Zhu†, Yuexin Ma†. **EvolvingGrasp: Evolutionary Grasp Generation via Efficient Preference Alignment.** (ICCV 2025)
3. **Yiming Zhong**, Yumeng Liu, Chuyang Xiao, Zemin Yang, Youzhuo Wang, Yufei Zhu, Yujing Sun, Xinge Zhu, Yuexin Ma†. **FreqPolicy: Frequency Autoregressive Visuomotor Policy with Continuous Tokens.** (NeurIPS 2025)
4. **Yiming Zhong***, Yaoyu He*, Zemin Yang*, Pengfei Tian, Yifan Huang, Qingqiu Huang, Xinge Zhu, Yuexin Ma†. **From Noise to Intent: Anchoring Generative VLA Policies with Residual Bridges.** (ICML 2026)
5. Heng Tao*, **Yiming Zhong***, Zemin Yang*, Yuexin Ma†. **FastGrasp: Learning-based Whole-body Control for Fast Dexterous Grasping with Mobile Manipulators.** (Under Review)
6. Yuhao Liu, **Yiming Zhong**†, Hanqing Wang, Shaocheng Yan, Yuhang Zhang, Wenzhou Lyu, Ziyang Ding, Wei Zhang, Xue Chao, Jin Pan, Yuexin Ma†, Xinge Zhu†. **UniAfford: Token-Routed Multitask Learning for Generalizable 2D-3D Affordance Perception.** (Under Review)
7. Linxi Chen*, Xiang Feng*, **Yiming Zhong***, Yiming Ren, Ziyi Wang, Zemin Yang, Yanming Zhu, Yuexin Ma. **One Framework, Multiple Deficiencies: Towards Unified Information-Deficient 3D Gaussian Reconstruction.** (Under Review)

Additional Publications

1. Youzhuo Wang*, Jiayi Ye*, Chuyang Xiao, **Yiming Zhong**, Heng Tao, Hang Yu, Yumeng Liu, Jingyi Yu, Yuexin Ma. **DexH2R: A Benchmark for Dynamic Dexterous Grasping in Human-to-Robot Handover.** (ICCV 2025)

2. Hanqing Wang*, Shaoyang Wang*, **Yiming Zhong**, Zemin Yang, Jiamin Wang, Zhiqing Cui, Jiahao Yuan, Yifan Han, Mingyu Liu, Yuexin Ma†. **Affordance-R1: Reinforcement Learning for Generalizable Affordance Reasoning in Multimodal Large Language Model.** (AAAI 2026 [Oral])
3. Zemin Yang, Yaoyu He, **Yiming Zhong**, Yuhao Zhang, Xinge Zhu, Yao Mu, Qingqiu Huang, Yuexin Ma†. **Implicit Drifting Policy: One-Step Action Generation via Conditional Expert Geometry.** (NeurIPS 2026)
4. Hanqing Wang, Mingyu Liu, Xiaoyu Chen, Chengwei Ma, **Yiming Zhong**, Wenti Yin, Yuhao Liu, Zhiqing Cui, Jiahao Yuan, Lu Dai, Zhiyuan Ma†, Hui Xiong†. **VideoAfford: Grounding 3D Affordance from Human-Object-Interaction Videos via Multimodal Large Language Model.** (NeurIPS 2026)
5. Xue Zhao, **Yiming Zhong**, Zemin Yang, Xiang Feng, Jin Pan, Xinbing Wang, Xinge Zhu, Yuexin Ma†, Nanyang Ye†. **LatentSightDrive: Progressive Foresight Internalization for Autonomous Driving.** (Under Review)

SELECTED RESEARCH EXPERIENCE

- **Uni-DGS: Unified Information-Deficient 3D Gaussian Reconstruction** Jan. 2026 – Sept. 2026
Under Review; Co-Project Lead; Advisor: Yuexin Ma
 - Co-led an undergraduate research project on unified 3D reconstruction under information-deficient observations, jointly defining the research direction, supervising weekly progress, and coordinating the overall project development.
 - Developed the core idea of combining scene-grounded internal reconstruction with external generative or restoration priors through a single-step diffusion fixer, enabling a shared framework for sparse-view reconstruction, 3D super-resolution, and deblurring.
 - Guided the project from problem formulation to paper completion, including method design, experimental planning, result analysis, and manuscript writing.
- **UniAfford: Token-Routed Multitask Learning for Generalizable 2D–3D Affordance Perception** Dec. 2025 – Apr. 2026
Under Review; Project Lead; Advisor: Yuexin Ma
 - Led an undergraduate research project on unified 2D–3D affordance perception, providing the core research direction, supervising weekly progress meetings, and coordinating model development and experimental validation.
 - Proposed the project-level design of a shared MLLM semantic space with token-based task routing and modality-specific 2D and 3D affordance decoders, enabling heterogeneous pixel-level and point-level supervision within a unified framework.
 - Directed the project from ideation to manuscript completion, including problem formulation, experimental design, result interpretation, and paper writing.
- **ResVLA: Anchoring Generative VLA Policies with Residual Bridges** Aug. 2025 – Feb. 2026
ICML 2026; Advisor: Yuexin Ma
 - Led the development of ResVLA, reframing generative visuomotor policy learning from “generation-from-noise” to “refinement-from-intent.”
 - Introduced spectral decomposition to separate low-frequency semantic intent from high-frequency action dynamics and modeled residual action refinement with diffusion bridges.
 - Demonstrated strong performance across diverse benchmarks, with improved robustness on long-horizon tasks and enhanced cross-embodiment generalization.
- **Fast Dexterous Grasping with Mobile Manipulators via Whole-body Control** Mar. 2025 – Feb. 2026
arXiv 2026; Advisor: Yuexin Ma
 - Co-led the development of FastGrasp, a learning-based framework that integrates grasp guidance, whole-body control, and tactile feedback for fast dexterous mobile manipulation.
 - Designed a two-stage reinforcement-learning pipeline that combines grasp proposal generation with execution-policy learning for stable and efficient grasping.
 - Enabled real-time coordination of the mobile base, robotic arm, and dexterous hand, with validation in both simulation and real-world deployment.

- **Frequency Autoregressive Visuomotor Policy with Continuous Tokens** *Mar. 2025 – Jun. 2025*
NeurIPS 2025; Advisor: Yuexin Ma
 - Led the exploration of frequency-domain action representations and hierarchical frequency modeling for structured robotic motion generation.
 - Introduced continuous latent action representations with diffusion-based decoding, avoiding discretization while preserving action continuity and autoregressive efficiency.
 - Achieved state-of-the-art performance on extensive robotic benchmarks in both task success and computational efficiency.
- **Evolutionary Grasp Generation via Efficient Preference Alignment** *Nov. 2024 – Mar. 2025*
ICCV 2025; Advisor: Yuexin Ma
 - Co-developed EvolvingGrasp, an evolutionary grasp-generation framework that iteratively refines grasp distributions for robust generalization in unstructured environments.
 - Introduced Handpose-wise Preference Optimization (HPO), formulating grasp adaptation as preference-guided posterior optimization.
 - Developed a Physics-Aware Consistency Model (PCM) to reduce sampling steps while preserving geometric and physical constraints.
- **Towards Universal Robotic Dexterous Grasping with Physics Awareness** *Jun. 2024 – Nov. 2024*
CVPR 2025 Highlight; Advisor: Yuexin Ma
 - Led the development of a physics-aware diffusion framework for dexterous grasp generation, integrating three physical constraints into both training and sampling.
 - Achieved state-of-the-art performance on five dexterous-grasping datasets, demonstrating strong cross-object generalization.
 - Introduced a large and diverse high-quality dexterous-grasping dataset to improve robustness across object categories.

HONORS AND AWARDS

National Scholarship	Oct. 2025
Huahong Scholarship	Nov. 2025
Outstanding Master Student	Nov. 2025
The International Mathematical Contest in Modeling (MCM), Finalist (Top 1%)	May 2023
China Undergraduate Mathematical Contest in Modeling (CUMCM), National First Prize (Top 0.5%)	Oct. 2022

ACADEMIC SERVICE AND EXPERIENCE

Conference Reviewer: CVPR, ICCV, ICLR, ICML, NeurIPS, IROS

Journal Reviewer: TVCG

Teaching: Teaching Assistant, CS100, ShanghaiTech University *Mar. 2025 – Jun. 2025*

Industry: Large-Model Algorithm Engineer Intern, Yinwang Intelligent Technologies *Aug. 2025 – Present*