Jingbo Zhang

Shenzhen, China | jbzhang6-c@my.cityu.edu.hk | (+86) 15727336166

Homepage | Google Scholar | Github

About Me

I am currently a senior researcher at Robotics X Lab, Tencent. I received my Ph.D. in Computer Science from City University of Hong Kong in 2025, under the supervision of Prof. Jing Liao. My research interests span machine learning, computer vision, computer graphics, large language models and robotics. Recently, my work has focused on 2D image editing; 3D/4D scene reconstruction, generation, and editing; as well as scene perception and interaction. I am also deeply interested in embodied intelligence and the development of world models that enable intelligent agents to navigate and interact effectively within complex environments.

Education

City University of Hong Kong, Department of Computer Science — Sep. 2020 – Oct. 2025

- Ph.D. in Computer Science and Technology, supervised by **Prof. Jing Liao**
- Research Interests: Computer Vision & Graphics; 3D Reconstruction, Generation, and Editing; Scene Perception & Interaction

Beihang University, School of Automation Science and Electrical Engineering — Sep. 2018 – Mar. 2020

- Ph.D. Candidate in Pattern Recognition and Intelligent System (Withdrew after the first year)

Beihang University, School of Aeronautic Science and Engineering — Sep. 2013 - Jul. 2018

- B.A. in Engineering Mechanics

Beihang University, School of Economics and Management — Sep. 2015 - Jul. 2018

- B.A. in Business Administration (Dual Degree)

Publications

- Jiaru Zhong, Yuxiang Yang, Wei Cui, Junliang Chen, **Jingbo Zhang**, Huaiyuan Xu, Jiahui Xu, Yijie Guo, Yi Wang, Qiang Zhang, Lap-Pui Chau. OE-VLN: Benchmarking Open-Ended Vision-Language Navigation under Natural Human Instructions. Under Review, 2026.
- Jiawei Hou, Shenghao Zhang, Can Wang, Zheng Gu, Yonggen Ling, Taiping Zeng, Xiangyang Xue, **Jingbo Zhang**. DetAny4D: Detect Anything 4D Temporally in a Streaming RGB Video[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2026.
- Zhe Wang, **Jingbo Zhang**, Wanchao Su, Tianyi Wei, Can Wang. WordCraft: Interactive Artistic Typography with Attention Awareness and Noise Blending. IEEE Transactions on Visualization and Computer Graphics. (Under Review)
- **Jingbo Zhang**, Xiaoyu Li, Hongliang Zhong, Qi Zhang, et al. HumanRef-GS: Image-to-3D Human Generation with Reference-Guided Diffusion and 3D Gaussian Splatting[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2025.
- Hongliang Zhong, Can Wang, **Jingbo Zhang**, Jing Liao. Generative Object Insertion in Gaussian Splatting with a Multi-View Diffusion Model[J]. Visual Informatics, 2025.
- Xiaoyu Li, Qi Zhang, Di Kang, Weihao Cheng, Yiming Gao, **Jingbo Zhang**, et al. Advances in 3D Generation: A Survey[J]. arXiv preprint arXiv:2401.17807, 2024.
- Ziyu Wan, **Jingbo Zhang**, Dongdong Chen, Jing Liao. High-Fidelity and Efficient Pluralistic Image Completion with Transformers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2024.
- **Jingbo Zhang**, Xiaoyu Li, Qi Zhang, et al. HumanRef: Single Image to 3D Human Generation via Reference-Guided Diffusion[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.
- **Jingbo Zhang**, Xiaoyu Li, Ziyu Wan, Can Wang, Jing Liao. Text2NeRF: Text-Driven 3D Scene Generation with Neural Radiance Fields[J]. IEEE Transactions on Visualization and Computer Graphics, 2024.

- Hongliang Zhong, **Jingbo Zhang**, Jing Liao. VQ-NeRF: Neural Reflectance Decomposition and Editing with Vector Quantization[J]. IEEE Transactions on Visualization and Computer Graphics, 2023.
- Ruixiang Jiang, Can Wang, **Jingbo Zhang**, et al. AvatarCraft: Transforming Text into Neural Human Avatars with Parameterized Shape and Pose Control[C]. IEEE/CVF International Conference on Computer Vision. 2023: 14371-14382.
- **Jingbo Zhang**, Xiaoyu Li, Ziyu Wan, Jing Liao. FDNeRF: Few-Shot Dynamic Neural Radiance Fields for Face Reconstruction and Expression Editing[C]. SIGGRAPH Asia 2022 Conference Papers. 2022: 1-9.
- **Jingbo Zhang**, Ziyu Wan, Jing Liao. Adaptive Joint Optimization for 3D Reconstruction with Differentiable Rendering[J]. IEEE Transactions on Visualization and Computer Graphics, 2022.
- Ziyu Wan, **Jingbo Zhang**, Dongdong Chen, Jing Liao. High-Fidelity Pluralistic Image Completion with Transformers[C]. IEEE/CVF International Conference on Computer Vision. 2021: 4692-4701.
- Yang Li, **Jingbo Zhang**, Weigang Cui, Heng Yuan, and Hualiang Wei. A Multiple Beta Wavelet-Based Locally Regularized Ultra-Orthogonal Forward Regression Algorithm for Time-Varying System Identification with Applications to EEG[J]. IEEE Transactions on Instrumentation and Measurement, 2019. ISSN 0018-9456.

Patents

- Haitao Lin, Jingshun Huang, Cheng Zhou, Yonggen Ling, He Zhang, Minglei Lu, Jingbo Zhang. 一种面向机器人操作的端到端开放词汇3D视觉语言模型(An End-to-End Open-Vocabulary 3D Vision-Language Model for Robotic Manipulation). CHN patent 2025120878CN. (Under Review)
- Jingbo Zhang, Jiawei Hou, Shenghao Zhang, Minglei Lu, Yonggen Ling, Haitao Lin, Yuzhen Liu. 一种面向视频输入的端到端开放集4D目标检测方法(An End-to-End Open-Set 4D Object Detection Method for Video Input). CHN patent 2025120130CN. (Under Review)
- Minglei Lu, Yonggen Ling, Shenghao Zhang, Yuzhen Liu, Haitao Lin, Jingbo Zhang. 一种基于全局多路优化的RGB视频6D位姿数据集标注方法(A Global Multi-Path Optimization-Based Annotation Method for RGB Video 6D Pose Datasets). CHN patent 2025120018CN. (Under Review)
- Minglei Lu, Junning Qiu, Yonggen Ling, Jingbo Zhang, Lingzhu Xiang. 一种基于动捕系统的机械臂运动学标定方法(A Kinematic Calibration Method for Robotic Arms Based on Motion Capture Systems). CHN patent 2025110345CN. (Under Review)
- Yuzhen Liu, Zhitong Huang, Jingbo Zhang, Zibo Zhang. 基于视觉语言模型引导的机器人探索策略(Vision-Language Model Guided Robotic Exploration Strategy). CHN patent 2025080364CN. (Under Review)
- Yang Li, Daxin Hao, Jingbo Zhang. 一种基于多小波基函数展开的锋电位时变格兰杰因果准确辨识方法(Accurate Identification of Time-Varying Granger Causality of Spike Potentials Based on Multi-Wavelet Basis Function Expansion). CHN patent CN108509933A[P].
- Yang Li, Jingbo Zhang, Weigang Cui, Song Xu, and Qinglei Hu. 一种基于beta小波基函数展开的时变非线性系统快速辨识方法(Fast Identification of Time-Varying Nonlinear Systems Based on Beta Wavelet Basis Function Expansion). CHN patent CN107967395A[P]

Research Experience

Embodied World Models Feb. 2026 - Present

- Developing action-conditioned world models to evaluate and train VLA policy prediction models.

3D scene perception and embodied navigation Apr. 2025 - Jan. 2026

- Constructed a hierarchical scene graph to serve as perceptual memory for an embodied robot.
- Conducted semantic recognition of objects and spatial perception in 3D environments.
- Developed and trained a large-scale model for human-robot interaction perception tasks.

3D Reconstruction, generation, and editing, image inpainting May. 2020 - Mar. 2025

- Reconstructed 3D models using multi-view RGB-D images, and jointly optimized their texture, geometry, and camera pose.
- Used few-shot dynamic frames to reconstruct an implicit 3D face model and perform novel view rendering and facial expression editing.
- Trained a NeRF model of 3D mixed materials based on multi-view RGB, and performed BRDF material

decomposition, editing and relighting on the model.

- Generated 3D scenes using prior of diffusion models.
- Generated 3D clothed humans from a single image based on pretrained diffusion models.

Object detection and image classification using RCNN series algorithms based on MS COCO, ImageNet and Pascal VOC datasets Sep. 2019 - Apr. 2020

- Compared the modeling ideas of RCNN, Fast-RCNN, Faster-RCNN, and Mask-RCNN, and conducted preliminary tests of the above models based on the MS-COCO 2014 and PASCAL-VOC 2007 databases.
- Tested the object detection accuracy of Faster-RCNN and Mask-RCNN with ResNet C4 and FPN (Feature Pyramid Network) and performed the object detection simulation using trained Faster-RCNN and Mask-RCNN models.
- Completed image classification tasks at Computer Vision Center of Tencent AI Lab.

Spiking Neural Networks for function connectivity analysis of hippocampal neural spikes Feb. 2019 - Jul. 2019

- Tested an ameliorated multiwavelet-based regularized forward orthogonal regression algorithm to improve the identification performance of a time-varying nonlinear generalized Laguerre-Volterra model, which is investigated for the nonstationary connectivity in spiking neural systems.
- Completed the simulation experiment and drafted the paper of the algorithm.

Signal processing and system identification for modeling scalp EEG data Nov. 2017 - Jan. 2019

- Proposed a novel parametric modeling algorithm to identify time-varying nonlinear systems, where a new class of multiple beta wavelet basis function is introduced to approximate time-varying coefficients of the nonstationary system.

Scholarship & Honors

Research Tuition Scholarship, CityUHK	2023
Outstanding Academic Performance Award, CityUHK	2022
Ph.D. Scholarship, CityUHK	2020 - 2024
Outstanding Graduate Thesis Award, BUAA	2018
National Encouragement Scholarship, BUAA	2017
Outstanding Academic Performance Scholarship, BUAA	2017
Model Student of Academic Records, BUAA	2017
Honorable Mention in the Zhou Peiyuan Mechanics Competition for College Students	2017
Honorable Mention in Interdisciplinary Contest in Modeling	2017
Scholarship for Excellent Social Work, BUAA	2016, 2017
Excellent Student-Cadre at BUAA	2016, 2017

Extra-curricular Activities

Research Intern, Embodied Intelligence Technology Center of Robotics X, Tencent	Sep. 2024 - Oct. 2025
Research Intern, Visual Computing Center of Tencent AI Lab	May. 2023 - Aug. 2024
Research Assistant, CityU Shenzhen Research Institute	May. 2020 - Aug. 2020
Research Intern, Computer Vision Center of Tencent AI Lab	Nov. 2019 - Apr. 2020
Volunteer service for the 2018 Beijing Marathon	Sep. 2018
Class Monitor	Sep. 2015 - Jun. 2018
Volunteer Teacher, Yunnan Mountainous Area	Jul. 2014 - Aug. 2014