Xiaolong ZHU

CONTACT INFORMATION

Parametrix.ai
2004, Bld 9B,
Shenzhenwan Science & Tech Ecological Garden
Nanshan, Shenzhen

Mobile: +86-14714930403
Skype: lucienzhu@hotmail.com
E-mail: lucienxlzhu@gmail.com
WWW: xiaolongzhu.org

RESEARCH INTERESTS

AI and Games, including *Bots*, *AI NPC*, *Generative AI*, *Agents*;
Computer Vision, including *Image Classification*, *Object Detection*, *Semantic Labelling*, *Landmark Localization*, *Neural Fields*;
Machine Learning, including *Random Forest*, *Support Vector Machines*, *Deep Learning*, *Reinforcement Learning*, *Diffusion Models*, *Foundation Models*;
Edge Computing, including *Image Processing*, *CPU/GPU/NPU Neural Network Inference*, *Heterogeneous Computing*;
Human-Computer Interaction, including *User Study*, *Prototyping*, *Gestural Interface*.

EDUCATION

The University of Hong Kong, Hong Kong SAR, China

Ph.D., *Computer Science*, **September 2010 - January 2016**

- Thesis Title: Hand Detection and Hand Shape and Posture Analysis in Images.
- Advisor: Dr. Kenneth K. Y. Wong

Peking University, Beijing, China

B.S., *Intelligence Science and Technology* **September 2006 - June 2010**

- Thesis Title: Segmentation and Classification of Range Image.
- *Excellent Undergraduate Thesis Award*.
- Advisor: Dr. Huijing Zhao

EXPERIENCE

Parametrix.ai, Tech Unicorn on Gaming AI Shenzhen, China

Vice President **Mar 2019 - Present**

- Business solutions and AI innovations, *e.g.*, Bots, NPCs and Generative AI;
- Platforms and algorithms design, *e.g.*, Distributed RL, Diffusion Models and Agents;
- Brands and UR, *e.g.*, NeuralMMO/Lux Competitions, MIT/THU/PKU Collaboration;
- Talent recruitment. Interviewed 500+ candidates of AI/Game/Frontend/Backend/Art/Design background;.

Tencent TEG, Shenzhen, China

Senior R&D Engineer, Tech Lead **Aug 2016 - Mar 2019**

- Led to deploy real-time face detection, landmarks and deformation on mobile phones;
- Developed and deployed real-time human pose estimation on iOS/Android phones;
- Helped to develop reinforcement learning for board game AI;
- Deployed real-time live video style transfer on iOS/Android phones;
- Developed an algorithm for real-time video style transfer;
- Developed several prototypes for AI Lab Vision Team.

R&D Engineer **Jul 2015 - Aug 2016**

- Implemented CTC model for end-to-end speech recognition, collaborating with WeChat Speech Team;
- Worked on prototyping news recommendation using DNN model;
- Implemented a prototype of service robot based on ROS/Turtlebot.

Lenovo IVC Lab, Hong Kong SAR, China

Research Intern **June 2013 - August 2013**

- Innovated new ways for image searching.
- Designed a prototype of touch-based image retrieval system and demonstrated it to CTO.

Microsoft Research Asia, Beijing, China

Research Intern **June 2012 - September 2012**

- Learned HCI workflow of problem solving;
- Designed visual feedback for in-air gesture recognition.

Youdao.com, Beijing, China

Software Engineer Intern **June 2010 - August 2010**

- Coded web front-end of a Location-based Social Network Service;
- Cooperated with web designer.

Peking University, Beijing, China

Undergraduate Research Assistant **September 2008 - June 2010**

- Participated in the *POSS* project, in 3D VCR Lab;
- Analyzed range data using computer vision methods.

TEACHING

Tsinghua Shenzhen International Graduate School, Shenzhen, China

Guest Teacher **September 2021 - Aug 2024**

- Guest speaker in Frontiers of AI Technology and Industrial Applications, 2022
- Guest speaker in Frontiers of AI Technology and Industrial Applications, 2021

The University of Hong Kong, Hong Kong SAR, China

Teaching Assistant **September 2010 - May 2014**

- Assisted Dr. Kenneth K.Y. Wong in Computer Vision;
- Assisted Dr. Kenneth K.Y. Wong in Computer Programming and Applications;
- Assisted Dr. Loretta Yi-King Choi in Topic in Computer Science: Visual Analysis.
- Assisted Dr. Kenneth K.Y. Wong in Computer Vision;
- Assisted Dr. Chun Kit Chui in Computer Programming and Applications;
- Assisted Dr. Kenneth K.Y. Wong in Computer Programming and Applications;

TALKS

- Algorithms and Applications of Massive Agents. Synced AI Annual Talks. in *Chinese*. 2023;
- How I teach AI to master the competition. 36Kr Documentary. in *Chinese*. 2022;
- Mobile AI Development on Arm Platform. Arm Developers Global Summit. in *Chinese*. 2018;
- Deploying AI on Mobile. Tencent HKU recruitment talk. 2018;
- Human Pose Estimation on Mobile. Tencent TLC. in *Chinese*. 2018;
- Panelist for LF DL session and Deep Learning Session. LC3 China. 2018;
- Learning Game of Go. Tencent AI Lab Academic Forum. in *Chinese*. 2018;

AWARDS

- Overseas High-Caliber Personnel (Level C) in Shenzhen, 2017-2023;
- Tencent Excellent R&D of the Year 2018, 2018;
- Tencent Technology Breakthrough of the Year 2017, 2017;
- Studentship of the University of Hong Kong, 2010-2014;
- Top 10 Undergraduate Thesis, School of EECS in Peking University, 2010;
- Wusi Scholarship in Peking University, 2009;
- Outstanding Volunteer in Beijing 2008 Olympic Games, 2008;
- First Class Honor in China Physics Olympic Games, Gansu, 2006.

TECHNICAL SKILLS
- Programming in: `Python`, `C/C++`, `Matlab`, `JavaScript/HTML/CSS`;
- Basic Experience in: `Objective-C`, `Processing`, `UNIX Shell scripting`;
- Native Mandarin speaker, fluent in English, very little Japanese and Cantonese;
- Operating Systems: Windows, Mac OS X.

SOCIAL ACTIVITIES
- TAC Member of LF Deep Learning Foundation, 2018-2019;
- Member of Tencent Open Source Working Group, 2018-2019;
- Co-founder of Tech Club of Tencent TEG, 2015-2016;
- Member of Information Technology Committee, The University of Hong Kong, 2012-2014;
- IT Officer of Postgraduate Association (PGSA) in The University of Hong Kong, 2011-2013;
- Volunteer as Media Assistant for Journalists in Games of the XXIX Olympiad, 2008.

HOBBIES

Board Games, Hiking, Indoor Climbing.

PUBLICATIONS

[Refereed Conference Papers]

17. Joseph Suarez, David Bloomin, Kyoung Whan Choe, Hao Xiang Li, Ryan Sullivan, Nishaanth Kanna, Daniel Scott, Rose Shuman, Herbie Bradley, Louis Castricato, Phillip Isola, Chenghui Yu, Yuhao Jiang, Qimai Li, Jiaxin Chen, **Xiaolong Zhu**. "Neural MMO 2.0: A Massively Multi-task Addition to Massively Multi-agent Learning." *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.

16. Kai Yang, Jian Tao, Jiafei Lyu, Chunjiang Ge, Jiaxin Chen, Qimai Li, Weihan Shen, **Xiaolong Zhu**, Xiu Li. "Using Human Feedback to Fine-tune Diffusion Models without Any Reward Model." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

15. Enhong Liu, Joseph Suarez, Chenhui You, Bo Wu, Bingcheng Chen, Jun Hu, Jiaxin Chen, **Xiaolong Zhu**, Clare Zhu, Julian Togelius, Sharada Mohanty, Weijun Hong, Rui Du, Yibing Zhang, Qinwen Wang, Xinhang Li, Zheng Yuan, Xiang Li, Yuejia Huang, Kun Zhang, Hanhui Yang, Shiqi Tang, Phillip Isola. "The NeurIPS 2022 Neural MMO Challenge: A Massively Multiagent Competition with Specialization and Trade." *NeurIPS 2022 Competitions Track*, 2023.

14. Yangkun Chen, Joseph Suarez, Junjie Zhang, Chenghui Yu, Bo Wu, Hanmo Chen, Hengman Zhu, Rui Du, Shanliang Qian, Shuai Liu, Weijun Hong, Jinke He, Yibing Zhang, Liang Zhao, Clare Zhu, Julian Togelius, Sharada Mohanty, Jiaxin Chen, Xiu Li, **Xiaolong Zhu**, Phillip Isola. "Benchmarking Robustness and Generalization in Multi-Agent Systems: A Case Study on Neural MMO." *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2023.

13. Hanmo Chen, Stone Tao, Jiaxin Chen, Weihan Shen, Xihui Li, Chenghui Yu, Sikai Cheng, **Xiaolong Zhu**, Xiu Li. "Emergent Collective Intelligence from Massive-agent Cooperation and Competition." *NeurIPS 2022 Deep RL Workshop*, 2023.

12. Yuhao Jiang, Kunjie Zhang, Qimai Li, Jiaxin Chen, **Xiaolong Zhu**. "Multi-Agent Path Finding via Tree LSTM." *AAAI 2023 MAPF Workshop*, 2022.

11. Rongqin Liang, Yuanheng Zhu, Zhentao Tang, Mu Yang, **Xiaolong Zhu**. "Proximal Policy Optimization with Elo-based Opponent Selection and Combination with Enhanced Rolling Horizon Evolution Algorithm." *IEEE Conference on Games (CoG)*, 2021.

10. Haozhi Huang, Hao Wang, Wenhan Luo, Lin Ma, Wenhao Jiang, **Xiaolong Zhu**, Zhifeng Li, and Wei Liu. "Real-Time Neural Style Transfer for Videos." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

9. **Xiaolong Zhu**, Wei Liu, Xuhui Jia and Kwan-Yee K. Wong. "A Two-Stage Detector for Hand Detection in Ego-Centric Videos." *Winter Conference on Applications of Computer Vision (WACV)*, 2016.

8. Xuhui Jia, Heng Yang, **Xiaolong Zhu**, Zhanghui Kuang, Yifeng Niu, Kwok-Ping Chan. "Reflective Regression of 2D-3D Face Shape Across Large Pose." *The British Machine Vision Conference (BMVC)*, 2016.

7. **Xiaolong Zhu**, Xuhui Jia and Kwan-Yee K. Wong. "Pixel-Level Hand Detection with Shape-aware Structured Forests." *Asian Conference on Computer Vision (ACCV)*, 2014.

6. **Xiaolong Zhu**, Ruoxin Sang, Xuhui Jia and Kwan-Yee K. Wong. "A Hand Shape Recognizer from Simple Sketches." *International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 2013.

5. Xuhui Jia, **Xiaolong Zhu**, Angran Lin and Kwok-Ping Chan. "Face Alignment using Structured Random Regressors Combined with Statistical Shape Model Fitting." *International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 2013.

4. **Xiaolong Zhu**, Kwan-Yee K. Wong. "Single-Frame Hand Gesture Recognition Using Color and Depth Kernel Descriptors." *IEEE International Conference on Pattern Recognition (ICPR)*, 2012.

3. Zhihu Chen, Kwan-Yee K. Wong, Yasuyuki Matsushita, **Xiaolong Zhu**, Miaomiao Liu. "Self-Calibrating Depth from Refraction." *IEEE International Conference on Computer Vision (ICCV)*, 2011.

2. **Xiaolong Zhu**, Huijing Zhao, Yiming Liu, Yipu Zhao, Hongbin Zha. "Segmentation and Classification of Range Image from an Intelligent Vehicle in Urban Environment." *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2010.

1. Huijing Zhao, Yiming Liu, **Xiaolong Zhu**, Yipu Zhao, Hongbin Zha. "Scene Understanding in a Large Dynamic Environment through a Laser-based Sensing." *IEEE International Conference on Robotics and Automation (ICRA)*, 2010.

[Journal Papers]

1. **Xiaolong Zhu**, Xuhui Jia, Kwan-Yee K. Wong. Structured Forests for Pixel-level Hand Detection and Hand Part Labelling. *Computer Vision and Image Understanding (CVIU)*, 2015.

2. Zhihu Chen, Kwan-Yee K. Wong, Yasuyuki Matsushita, **Xiaolong Zhu**. Depth from Refraction Using a Transparent Medium with Unknown Pose and Refractive Index. *International Journal of Computer Vision (IJCV)*, 2012.

PATENTS

[In English]

1. Method and apparatus for training gaze tracking model, and method and apparatus for gaze tracking. US-11797084-B2

2. Facial expression synthesis method and apparatus, electronic device, and storage medium. US-11030439-B2

3. Method and apparatus for training neural network model used for image processing, and storage medium. US-2019228264-A1

4. Method and apparatus for recognizing postures of multiple persons, electronic device, and storage medium. EP-3876140-B1

5. Method and apparatus for training pose recognition model, and method and apparatus for image recognition. US-11907848-B2
6. Augmented reality processing method, object recognition method, and related apparatus. EP-3617995-A1
7. Story monitoring method when robot takes elevator, electronic device, and computer storage medium. US-11242219-B2
8. Method and apparatus for generating music. US-11301641-B2
9. Action recognition method and apparatus, and human-machine interaction method and apparatus. US-11710351-B2
10. Camera orientation tracking method and apparatus, device, and system. EP-3798983-B1
11. Neural network model deployment method, prediction method, and apparatus. EP-3614316-A1
12. Video data processing method and apparatus, and storage medium. US-11461876-B2
13. Video image processing method and apparatus. US-10880458-B2
14. Image recognition method and apparatus, electronic device, and readable storage medium using an update on body extraction parameter and alignment parameter. US-11417095-B2
15. Foreground data generation method and method for applying same, related apparatus, and system. US-2021279888-A1
16. Image processing method and apparatus. US-11200680-B2
17. Method and apparatus for training gaze tracking model, and method and apparatus for gaze tracking. US-11797084-B2
18. Facial expression synthesis method and apparatus, electronic device, and storage medium. US-11030439-B2
19. Method and apparatus for training neural network model used for image processing, and storage medium. US-2019228264-A1
20. Method and apparatus for recognizing postures of multiple persons, electronic device, and storage medium. EP-3876140-B1
21. Method and apparatus for training pose recognition model, and method and apparatus for image recognition. US-11907848-B2
22. Augmented reality processing method, object recognition method, and related apparatus. EP-3617995-A1
23. Story monitoring method when robot takes elevator, electronic device, and computer storage medium. US-11242219-B2
24. Method and apparatus for generating music. US-11301641-B2
25. Action recognition method and apparatus, and human-machine interaction method and apparatus. US-11710351-B2
26. Camera orientation tracking method and apparatus, device, and system. EP-3798983-B1
27. Neural network model deployment method, prediction method, and apparatus. EP-3614316-A1

28. Video data processing method and apparatus, and storage medium. US-11461876-B2
29. Video image processing method and apparatus. US-10880458-B2
30. Image recognition method and apparatus, electronic device, and readable storage medium using an update on body extraction parameter and alignment parameter. US-11417095-B2
31. Foreground data generation method and method for applying same, related apparatus, and system. US-2021279888-A1
32. Image processing method and apparatus. US-11200680-B2
33. Control method and device for interactive task, storage medium and computer equipment. CN-110639208-B
34. Model training method, model calling equipment and readable storage medium. CN-110782004-B
35. Model training method, model using method, computer device, and storage medium. CN-111841018-B
36. Interaction model training method, device, computer equipment and storage medium. CN-113509726-B
37. Data sampling method and device, computer equipment and storage medium. CN-113032621-A
38. Intelligent agent control method and device, computer equipment and storage medium. CN-112905013-A
39. Level setting method and device, computer equipment and storage medium. CN-113134238-A
40. Modeling method and device of intelligent agent, equipment and storage medium. CN-115645910-A
41. Game behavior planning method, device, equipment and storage medium. CN-115624759-A
42. Model training method, action strategy making method, server and storage medium. CN-115944924-A
43. Video frame data sampling method and device, computer equipment and storage medium. CN-112925949-A
44. Virtual object control method, virtual object control device, virtual object model training method, virtual object model training device and computer equipment. CN-112933605-A
45. Intelligent agent training method, computer equipment and storage medium. CN-115759284-A
46. Feature analysis method, device, computer equipment and storage medium. CN-113139447-B