

Zhaochong An

Email	anzhaochong@outlook.com	Homepage	zhaochongan.github.io
GitHub	github.com/ZhaochongAn	Scholar	scholar.google.com
LinkedIn	linkedin.com/in/zhaochong-an	Location	Copenhagen, Denmark

Research Interests

Generative Modelling: developing novel algorithms for controllable and coherent long-form video generation, as well as post-training methods for reward-aligned generative models.

Multimodal Foundation Models: designing novel architectures and data curation pipelines to enhance large multimodal models (LMMs) for unified any-to-any generation and multimodal understanding.

Education

Oct 2023 – Present	Ph.D. in Computer Science – University of Copenhagen Advisor: Prof. Serge Belongie Affiliated with the Pioneer Centre for Artificial Intelligence; ELLIS PhD program Expected graduation: 2027
Sep 2021 – Sep 2023	M.Sc. in Computer Science – ETH Zurich GPA: 5.52/6.0; Advisor: Prof. Luc Van Gool (Computer Vision Lab) Major: Machine Intelligence; Minor: Data Management Systems
Aug 2017 – Jul 2021	B.Eng. in Computer Science and Technology – Harbin Institute of Technology GPA: 92.78/100

Publications

*Author names ordered as published; * denotes equal contribution.*

MODUS: Decoder-only Any-to-Any Modeling of Diverse Modalities

Mingqiao Ye*, **Zhaochong An***, Zhitong Gao, Xian Liu, Oğuzhan Fatih Kar, Jesse Allardice, Roman Bachmann, David Mizrahi, François Fleuret, Chuan Li, Amir Zadeh, Serge Belongie, Afshin Dehghan, Amir Zamir
*International Conference on Machine Learning (ICML), 2026. *Equal contribution.*

VGGRPO: Towards World-Consistent Video Generation with 4D Latent Reward

Zhaochong An, Orest Kupyn, Théo Uscidda, Andrea Colaco, Karan Ahuja, Serge Belongie, Mar Gonzalez-Franco, Marta Tintore Gazulla
European Conference on Computer Vision (ECCV), 2026. Work done at Google.

OneStory: Coherent Multi-Shot Video Generation with Adaptive Memory

Zhaochong An, Menglin Jia, Haonan Qiu, Zijian Zhou, Xiaoke Huang, Zhiheng Liu, Weiming Ren, Kumara Kahatapitiya, Ding Liu, Sen He, Chenyang Zhang, Tao Xiang, Fanny Yang, Serge Belongie, Tian Xie
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2026. Work done at Meta.

TUNA: Taming Unified Visual Representations for Native Unified Multimodal Models

Zhiheng Liu, Weiming Ren, Haozhe Liu, Zijian Zhou, Shoufa Chen, Haonan Qiu, Xiaoke Huang, **Zhaochong An**, Fanny Yang, Aditya Patel, Viktar Atliha, Tony Ng, Xiao Han, Chuyan Zhu, Chenyang Zhang, Ding Liu, Juan-Manuel Perez-Rua, Sen He, Jürgen Schmidhuber, Wenhui Chen, Ping Luo, Wei Liu, Tao Xiang, Jonas Schult, Yuren Cong
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2026. Highlight.

Scaling Zero-Shot Reference-to-Video Generation

Zijian Zhou, Shikun Liu, Haozhe Liu, Haonan Qiu, **Zhaochong An**, Weiming Ren, Zhiheng Liu, Xiaoke Huang, Kam Woh Ng, Tian Xie, Xiao Han, Yuren Cong, Hang Li, Chuyan Zhu, Aditya Patel, Tao Xiang, Sen He
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2026.

HiStream: Efficient High-Resolution Video Generation via Redundancy-Eliminated Streaming

Haonan Qiu, Shikun Liu, Zijian Zhou, **Zhaochong An**, Weiming Ren, Zhiheng Liu, Jonas Schult, Sen He, Shoufa Chen, Yuren Cong, Tao Xiang, Ziwei Liu, Juan-Manuel Perez-Rua
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Findings, 2026.

VecGlypher: Unified Vector Glyph Generation with Language Models

Xiaoke Huang, Bhavul Gauri, Kam Woh Ng, Tony Ng, Mengmeng Xu, Zhiheng Liu, Weiming Ren, **Zhaochong An**, Zijian Zhou, Haonan Qiu, Yuyin Zhou, Sen He, Ziheng Wang, Tao Xiang, Xiao Han

IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2026.

Multimodality Helps Few-shot 3D Point Cloud Semantic Segmentation

Zhaochong An, Guolei Sun, Yun Liu, Runjia Li, Min Wu, Ming-Ming Cheng, Ender Konukoglu, Serge Belongie
*International Conference on Learning Representations (ICLR), 2025. **Spotlight – Top 5.1% of submissions.***

Generalized Few-shot 3D Point Cloud Segmentation with Vision-Language Model

Zhaochong An, Guolei Sun, Yun Liu, Runjia Li, Junlin Han, Ender Konukoglu, Serge Belongie
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025.

What You Have is What You Track: Adaptive and Robust Multimodal Tracking

Yuedong Tan, Jiawei Shao, Eduard Zamfir, Ruanjun Li, **Zhaochong An**, Chao Ma, Danda Paudel, Luc Van Gool, Radu Timofte, Zongwei Wu
International Conference on Computer Vision (ICCV), 2025.

DreamBeast: Distilling 3D Fantastical Animals with Part-Aware Knowledge Transfer

Runjia Li, Junlin Han, Luke Melas-Kyriazi, Chunyi Sun, **Zhaochong An**, Zhongrui Gui, Shuyang Sun, Philip Torr, Tomas Jakab
International Conference on 3D Vision (3DV), 2025.

Rethinking Few-shot 3D Point Cloud Semantic Segmentation

Zhaochong An, Guolei Sun, Yun Liu, Fayao Liu, Zongwei Wu, Dan Wang, Luc Van Gool, Serge Belongie
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.

Experiences

Apr 2026 –
Present

Meta

Research Scientist Intern

- Research on the pretraining and generalization of world action models (ongoing).

Oct 2025 –
Feb 2026

Google

Research Scientist Intern

- Developed VGGRPO for world-consistent video generation via a 4D latent reward; published as first author at ECCV 2026.

Jun 2025 –
Sep 2025

Meta

Research Scientist Intern

- Developed OneStory, a multi-shot video generator that autoregressively synthesises shots with an adaptive memory mechanism; published as first author at CVPR 2026.
- Contributed to unified multimodal modelling (TUNA), reference-to-video generation (Saber), and other efforts accepted to CVPR 2026.

Invited Talks

Mar 2026

World-Consistent Video Generation with 4D Latent Reward

Google

Jul 2025

Few-shot 3D Understanding in the Era of Foundation Models

INSAIT (Institute for Computer Science, Artificial Intelligence and Technology)

Academic Services

Reviewer: CVPR, NeurIPS, ICML, ICLR, ECCV, ICCV, ACM MM.

Skills

Programming: Python (fluent), C/C++ , Java, Scala, MATLAB, HTML

Machine Learning: PyTorch, TensorFlow, NumPy, Pandas

Languages: Mandarin (native), English (professional working proficiency)

Standardized Tests: TOEFL 111 (R30 / L30 / S23 / W28); GRE (V158 / Q170 / AWA 4.0)