Visual Geometry Group (VGG)
Department of Engineering Science
University of Oxford
✉ *cxzheng@robots.ox.ac.uk*
chuanxiaz.com
lyndonzheng

Dr. Chuanxia Zheng

Research Interests

His research interests focus on computer vision and machine learning, especially for creative AI. He has done a wide range of work on 2D and 3D scene synthesis, with the goal of *synthesizing a photorealistic physical world* via generative AI. In particular, on topics:

- 3D geometry and appearance reconstruction from images or videos.
- 3D editing via object-centric perception.
- Generative models for physical world perception and reasoning.
- Multi-modalities (1D, 2D, 3D, and 4D) generation and understanding.

Professional experience

2024– **Marie Skłodowska-Curie Actions (MSCA) Fellow**, *University of Oxford*, UK.
3D scene creation from images or videos

2022–24 **Postdoctoral Research Fellow**, *University of Oxford*, UK.
2D and 3D scene synthesis

2021–22 **Research Fellow**, *Monash University*, Australia.
Codebook learning for 2D and 3D synthesis

Education

2017–21 **Doctor of Philosophy (PhD)**.
Nanyang Technological University
School of Computer Science and Engineering, Singapore
Thesis: *Synthesizing Photorealistic Images with Deep Generative Learning*
Outstanding PhD Thesis Award, Advisors: Tat-Jen Cham and Jianfei Cai

2014–17 **Master of Science (MSc) in computer science**.
Beihang University, Beijing, China
Thesis: *Context-based Indoor Scene Understanding for Mobile Robot*
Advisors: Jianhua Wang and Weihai Chen

2010–14 **Bachelor of Science in information engineering**.
Beijing Jiaotong University, Beijing, China
Thesis: *Image Retrieval based on Visual Saliency*
Highest Honours (Outstanding Graduate of Beijing), Advisor: Ze Liu

Research Experience

2022- **Research Fellow**, *University of Oxford*, UK, Prof. Andrea Vedaldi.
Research interests: 3D reconstruction from limited images or videos

2021-22 **Research Fellow**, *Monash University*, Australia, Prof. Jianfei cai.
Research interests: nature scene generation and completion

2017-21 **PhD**, *Nanyang Technological University*, Singapore, Prof. Nadia Thalmann.
Research interests: photorealistic image generation

Grants

2024- €236,748, PI, "Synthesizing Photorealistic 3D Scene", HORIZON-MSCA

2024- €5,910, Co-PI, "Object-Centric 3D Reconstruction and Decomposition", Bavarian Funding.

Awards and other recognitions

2024 HORIZON Marie Skłodowska-Curie (HORIZON-MSCA) Fellowship
2023 Outstanding Reviewer Award, CVPR
2022 Scholar/Travel Award, NeurIPS
2022 Wallenberg-NTU PPF, NTU, Singapore
2022 Outstanding PhD Thesis Award, NTU, Singapore
2021 Outstanding Reviewer Award, IEEE Transactions on Multimedia (TMM)
2017 NTU Research Scholarship
2014 Outstanding Graduate of Beijing
2012 Hanergy Scholarship Award (**Top 1%**)
2011 Siemens Scholarship Award (**Top 1%**)

Press Coverage

2023 Sber.ru: MoVQ — 0.1 means a lot for text-image generation Kandinsky 2.1 (Github: 2.7K)
2022 Phys.org: Researchers unravel cell biology through artificial intelligence
2022 NTU News: NTU SCSE Outstanding PhD Thesis Award 2022
2022 Zhuanzhi: How to create photorealistic images? Ph.D. Thesis by Dr. Zheng
2021 kknews, Sohu, NetEase: AgileGAN — a tool for creating stylized portraits (Demo: 10K/week)

Service to the academic community

- **Area Chair.** ACM Multimedia 2024, BMVC 2024.
- **Reviewer for international journals.** IEEE Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**), International Journal on Computer Vision (**IJCV**), IEEE Transactions on Image Processing (**TIP**), IEEE Transactions on Multimedia(**TMM**), Computer Vision and Image Understanding (**CVIU**), The Visual Computer (**TVC**).
- **Reviewer for international conferences.** IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**) 2020-2024, European Conference on Computer Vision (**ECCV**) 2020, 2022, 2024, International Conference on Computer Vision (**ICCV**) 2019, 2021, 2023, International Conference on Neural Information Processing Systems (**NeurIPS**) 2022-2024, International Conference on Learning Representations (**ICLR**) 2021-2024, International Conference on Machine Learning (**ICML**) 2023, International Conference on Computer Graphics (**SIGGRAPH**) 2021,2022, International Conference on Robotics and Automation (**ICRA**) 2023.

International workshops

2024 "Second Workshop for Learning 3D with Multi-View Supervision" at the IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**) with Abdullah Hamdi, Silvio Giancola, Guocheng Qian, Jinjie Mai, Sara Rojas Martinez, Bernard S. Ghanem, and Yash Bhalgat.

Mentoring and student supervision

PhD

2023- Brandon Smart, University of Oxford, co-supervised with Prof. Victor Prisacariu
2023- Ruining Li, University of Oxford, co-supervised with Prof. Andrea Vedaldi
2023- Tianhao Wu, NTU, co-supervised with Prof. Tat-Jen Cham
2022-23 Minghui Hu, NTU, three terms with Prof. Tat-Jen Cham
2021- Yuedong Chen, Monash University, co-supervised with Prof. Jianfei Cai

Master

2024- Michael Neumayr, TUM, co-supervised with Dr. Yan Xia and Prof. Daniel Cremers

2024- Filip Skubacz, TUM, co-supervised with Dr. Yan Xia and Prof. Daniel Cremers

2024- Nina Kirakosyan, TUM, co-supervised with Dr. Yan Xia and Prof. Daniel Cremers

2024- Wenbo Ji, TUM, co-supervised with Dr. Yan Xia and Prof. Daniel Cremers

Teaching

2023-23 **Teaching Assistant**, *B16: Software Engineering*, Undergraduate, University of Oxford.

2023-23 **Teaching**, *Generative AI*, Graduate, University of Oxford.

2018–20 **Teaching Assistant**, *Advanced Digital Image Processing*, Graduate, NTU.

2018–20 **Teaching Assistant**, *Human-Computer Interaction*, Undergraduate, NTU.

2018–19 **Teaching Assistant**, *Engineering Mathematics*, Undergraduate, NTU.

Invited talks

2023 **Visiting the Invisible via Generative AI**, *University of Science and Technology*, China.

2023 **Codebook Leaning for Generative AI**, *Harbin Institute of Technology*, China.

2023 **Codebook Leaning for Generative AI**, *Nanyang Technological University*, Singapore, NTU.

2023 **Codebook Leaning for Generative AI**, *University of Oxford*, UK.

2022 **Synthesizing Photorealistic Scenes**, *Nanyang Technological University*, Singapore, Link.

2022 **Synthesizing Photorealistic Scenes**, *University of Oxford*, UK.

2022 **Synthesizing Photorealistic Scenes**, *ETH*, Zürich.

2022 **Synthesizing Photorealistic Scenes**, *University of Science and Technology*, China.

2019 **Pluralistic Image Completion**, *Nanyang Technological University*, Singapore.

2018 **Depth Estimation from Single 2D Image**, *Nanyang Technological University*, Singapore.

Publications

[26] **Chuanxia Zheng**, Guoxian Song, Tat-Jen Cham, Jianfei Cai, Linjie Luo, and Dinh Phung. Bridging global context interactions for high-fidelity pluralistic image completion. ***TPAMI***, 2024. URL: https://chuanxiaz.com/picformer/.

[25] **Chuanxia Zheng** and Andrea Vedaldi. Free3d: Consistent novel view synthesis without 3d representation. In ***CVPR***, 2024. URL: https://chuanxiaz.com/free3d/.

[24] Guanqi Zhan, **Chuanxia Zheng**, Weidi Xie, and Andrew Zisserman. Amodal ground truth and completion in the wild. In ***CVPR***, 2024. URL: https://www.robots.ox.ac.uk/ vgg/research/.

[23] Minghui Hu, Jianbin Zheng, **Chuanxia Zheng**, Chaoyue Wang, Dacheng Tao, and Tat-Jen Cham. One more step: A versatile plug-and-play module for rectifying diffusion schedule flaws and enhancing low-frequency controls. In ***CVPR***, 2024. URL: https://jabir-zheng.github.io/OneMoreStep/.

[22] Tianhao Wu, **Chuanxia Zheng**, and Tat-Jen Cham. Panodiffusion: 360-degree panorama outpainting via diffusion. In ***ICLR***, 2024. URL: https://sm0kywu.github.io/panodiffusion/.

[21] Minghui Hu, Jianbin Zheng, Daqing Liu, **Chuanxia Zheng**, Chaoyue Wang, Dacheng Tao, and Tat-Jen Cham. Cocktail: Mixing multi-modality control for text-conditional image generation. In ***NeurIPS***, 2023. URL: https://mhh0318.github.io/cocktail/.

[20] **Chuanxia Zheng** and Andrea Vedaldi. Online clustered codebook. In ***ICCV***, 2023. URL: https://chuanxiaz.com/cvq/.

[19] Long Tung Vuong, Trung Le, He Zhao, **Chuanxia Zheng**, Mehrtash Harandi, Jianfei Cai, and Dinh Phung. Vector quantized wasserstein auto-encoder. In ***ICML***, 2023.

[18] Minghui Hu, **Chuanxia Zheng**, Heliang Zheng, Tat-Jen Cham, Zuopeng Yang, Chaoyue Wang, Dacheng Tao, and Ponnuthurai N. Suganthan. Unified discrete diffusion for simultaneous vision-language generation. In ***ICLR***, 2023. URL: https://mhh0318.github.io/unid3/.

[17] **Chuanxia Zheng**, Tung Vuong, Jianfei Cai, and Dinh Phung. Movq: Modulating quantized vectors for high-fidelity image generation. In ***NeurIPS***, 2022. URL: https://chuanxiaz.com.

[16] Jyothsna Vasudevan*, **Chuanxia Zheng***, James G. Wan, Tat-Jen Cham, Lim Chwee Teck, and Javier G. Fernandez. From qualitative data to correlation using deep generative networks: Demonstrating the relation of nuclear position with the arrangement of actin filaments. ***PloS one***, 17(7):e0271056, 2022.

[15] Qianyi Wu, Xian Liu, Yuedong Chen, Kejie Li, **Chuanxia Zheng**, Jianfei Cai, and Jianming Zheng. Object-compositional neural implicit surfaces. In ***ECCV***, 2022. URL: https://wuqianyi.top/objectsdf/.

[14] Yuedong Chen, Qianyi Wu, **Chuanxia Zheng**, Tat-Jen Cham, and Jianfei Cai. Sem2nerf: Converting single-view semantic masks to neural radiance fields. In ***ECCV***, 2022. URL: https://donydchen.github.io/sem2nerf/.

[13] **Chuanxia Zheng**, Tat-Jen Cham, Jianfei Cai, and Dinh Phung. Bridging global context interactions for high-fidelity image completion. In ***CVPR***, pages 11512–11522, June 2022. URL: https://chuanxiaz.com/tfill/.

[12] **Chuanxia Zheng**, Duy-Son Dao, Guoxian Song, Tat-Jen Cham, and Jianfei Cai. Visiting the invisible: Layer-by-layer completed scene decomposition. *International Journal of Computer Vision (**IJCV**)*, 129(12):3195–3215, 2021. URL: https://chuanxiaz.com/vinv/.

[11] Yujun Cai, Yiwei Wang, Yiheng Zhu, Tat-Jen Cham, Jianfei Cai, Junsong Yuan, Jun Liu, **Chuanxia Zheng**, Sijie Yan, Henghui Ding, Xiaohui Shen, Ding Liu, and Nadia Magnenat Thalmann. A unified 3d human motion synthesis model via conditional variational auto-encoder. In ***ICCV***, pages 11645–11655, 2021.

[10] **Chuanxia Zheng**, Tat-Jen Cham, and Jianfei Cai. Pluralistic free-form image completion. *International Journal of Computer Vision (**IJCV**)*, 129(10):2786–2805, 2021. URL: https://chuanxiaz.com/pic/.

[9] Guoxian Song, Linjie Luo, Jing Liu, Wan-Chun Ma, Chunpong Lai, **Chuanxia Zheng**, and Tat-Jen Cham. Agilegan: stylizing portraits by inversion-consistent transfer learning. *ACM Transactions on Graphics (**TOG**)*, 40(4):1–13, 2021. URL: https://guoxiansong.github.io.

[8] **Chuanxia Zheng**, Tat-Jen Cham, and Jianfei Cai. The spatially-correlative loss for various image translation tasks. In ***CVPR***, 2021. URL:https://chuanxiaz.com/flsesim/.

[7] **Chuanxia Zheng**, Tat-Jen Cham, and Jianfei Cai. Pluralistic image completion. In ***CVPR***, pages 1438–1447, 2019. URL: https://chuanxiaz.com/pic/.

[6] Tianyi Zhang, Jingyi Yang, **Chuanxia Zheng**, Guosheng Lin, Jianfei Cai, and Alex C Kot. Task-in-all domain adaptation for semantic segmentation. In ***VCIP***, pages 1–4, 2019.

[5] **Chuanxia Zheng**, Tat-Jen Cham, and Jianfei Cai. T2net: Synthetic-to-realistic translation for solving single-image depth estimation tasks. In ***ECCV***, pages 767–783, 2018. URL: https://chuanxiaz.com/synthetic2real/.

[4] **Chuanxia Zheng**, Jianhua Wang, Weihai Chen, and Xingming Wu. Multi-class indoor semantic segmentation with deep structured model. ***TVCJ***, 34(5):735–747, 2018.

[3] Jianhua Wang, **Chuanxia Zheng**, Weihai Chen, and Xingming Wu. Learning aggregated features and optimizing model for semantic labeling. ***TVCJ***, 33(12):1587–1600, 2017.

[2] **Chuanxia Zheng**, Jianhua Wang, Weihai Chen, and Xingming Wu. Semantic segmentation based on aggregated features and contextual information. In ***ROBIO***. IEEE, 2016.

[1] Jianhua Wang, **Chuanxia Zheng**, Weihai Chen, and Xingming Wu. Learning contextual information for indoor semantic segmentation. In ***ICIEA***, pages 1639–1644. IEEE, 2016.

Preprint

[9] Stanislaw Szymanowicz, Eldar Insafutdinov, **Chuanxia Zheng**, Dylan Campbell, Jao F. Henriques, Christian Rupprecht, and Andrea Vedaldi. Flash3d: Feed-forward generalisable scene reconstruction from a single image. *Under reviewer*.

[8] Yuedong Chen, Haofei Xu, **Chuanxia Zheng**, Bohan Zhuang, Andrea Vedaldi, Tat-Jen Cham, and Jianfei Cai. Mvsplat360: Benchmarking 360-degree generalizable 3d novel view synthesis from sparse views. *Under reviewer*.

[7] Yuedong Chen, Haofei Xu, **Chuanxia Zheng**, Bohan Zhuang, Marc Pollefeys, Andreas Geiger, Tat-Jen Cham, and Jianfei Cai. Mvsplat: Efficient 3d gaussian splatting from sparse multi-view images. *Under reviewer*.

[6] Ruining Li, **Chuanxia Zheng**, Christian Rupprecht, and Andrea Vedaldi. Dragapart: Learning a part-level motion prior for articulated objects. *Under reviewer*.

[5] Tianhao Wu, **Chuanxia Zheng**, Tat-Jen Cham, and Qianyi Wu. Clusteringsdf: Self-organized neural implicit surfaces for 3d decomposition. *Under reviewer*.

[4] Guanqi Zhan, **Chuanxia Zheng**, Weidi Xie, and Andrew Zisserman. What does stable diffusion know about the 3d scene. *Under reviewer*.

[3] LongTung Vuong, **Chuanxia Zheng**, Manh Luong, Thanh-Toan Do, Dinh Phung, and Trung Le. Kefi: Kernel-based feature identification for generalizable classification. *Under reviewer*.

[2] Yuedong Chen, Haofei Xu, Qianyi Wu, **Chuanxia Zheng**, Tat-Jen Cham, and Jianfei Cai. Explicit correspondence matching for generalizable neural radiance fields. *Under reviewer*.

[1] Yuzhu Ji, **Chuanxia Zheng**, and Tat-Jen Cham. One-shot human motion transfer via occlusion-robust flow prediction and neural texturing. *Under reviewer on* ***TNNLS***.