Shengqu Cai

shengqu@cs.stanford.edu | https://primecai.github.io

Education

since 2023 | PhD in Computer Science, **Stanford University**, United States
Advised by Prof. Gordon Wetzstein and Prof. Leonidas Guibas. GPA: 4.3/4.3

2020 - 2023 | MSc in Computer Science, **ETH Zürich**, Switzerland
Advised by Prof. Luc Van Gool. GPA: 5.7/6.0 (Major GPA: 6.0/6.0)

2017 - 2020 | BSc (Hons) in Computer Science, **King's College London**, United Kingdom
First Honour. Average: 90% (GPA: 4.0/4.0, ≈ top 1%)

Industrial Experience

2025 - 2026 | Research Intern at **NVIDIA**, Santa Clara, USA
Advised by Arash Vahdat.

2025 | Research Intern at **ByteDance Seed**, San Jose, USA
Advised by Ceyuan Yang.

2023 | Research Intern at **Adobe Research**, California, USA
Advised by Duygu Ceylan.

2021 | Research Student at **ETH Zürich CVL & Toyota TRACE**, Zürich, Switzerland
Advised by Anton Obukhov.

Publication

* and † indicate equal contribution.

[21] Mixture of Contexts for Long Video Generation,
in **ICLR** 2026.
Shengqu Cai, Ceyuan Yang, Lvmin Zhang, Yuwei Guo, Junfei Xiao,
Ziyan Yang, Yinghao Xu, Zhenheng Yang, Alan Yuille, Leonidas Guibas,
Maneesh Agrawala, Lu Jiang, Gordon Wetzstein

[20] pi-H3: Video Generation in 2–4 Steps Without a VAE Decoder,
Blog post, 2026.
Hansheng Chen, **Shengqu Cai**, Jiaming Song, Gordon Wetzstein,
Xinchen Yan, Leonidas Guibas

[19] FramePack: Frame Context Packing and Drift Prevention in
Next-Frame-Prediction Video Diffusion Models,
in **NeurIPS** 2025 (Spotlight).
Lvmin Zhang, **Shengqu Cai**, Muyang Li, Gordon Wetzstein, Maneesh Agrawala

[18] Mode Seeking meets Mean Seeking for Fast Long Video Generation,
in **ICML** 2026.
Shengqu Cai, Weili Nie*, Chao Liu*, Julius Berner, Lvmin Zhang, Nanye Ma,
Hansheng Chen, Maneesh Agrawala, Leonidas Guibas, Gordon Wetzstein, Arash Vahdat

[17] Decoding Without a Decoder: Direct Pixel Prediction from
Diffusion Transformers,
Hansheng Chen, **Shengqu Cai**, Jan Ackermann, Rayhan Zirvi,
Gordon Wetzstein, Xinchen Yan, Leonidas Guibas

[16] TinyHistory: Lightweight Video History Embeddings via Two-Stage Context Learning,
in **ECCV** 2026.
Lvmin Zhang, **Shengqu Cai**, Muyang Li, Chong Zeng, Beijia Lu,
Anyi Rao, Song Han, Gordon Wetzstein, Maneesh Agrawala

[15] GeoFlow: Enforcing Implicit Geometric Consistency in Video Generation,
in **arXiv** 2026.
Jan Ackermann, **Shengqu Cai***, Boyang Deng*, Zhengfei Kuang*,
Songyou Peng, Gordon Wetzstein

[14] Effective Multi-sensor Conditioning for Street-view Novel-view Synthesis,
in **NeurIPS** 2026.
Zhengfei Kuang, Adam Sun, Liyuan Zhu, Tong Wu, **Shengqu Cai**,
Jonathan Tremblay, Iro Armeni, Ehsan Adeli, Lior Yariv, Gordon Wetzstein

[13] Generated Reality: Human-centric World Simulation using Interactive
Video Generation with Hand and Camera Control,
in **CVPR Findings** 2026.
Linxi Xie*, Lisong C. Sun*, Ashley Neall, Tong Wu, **Shengqu Cai**,
Gordon Wetzstein

[12] BulletTime: Decoupled Control of Time and Camera Pose for Video Generation,
in **CVPR** 2026.
Yiming Wang, Qihang Zhang*, **Shengqu Cai***, Tong Wu†, Jan Ackermann†,
Zhengfei Kuang†, Yang Zheng†, Frano Rajic†, Siyu Tang, Gordon Wetzstein

[11] Captain Cinema: Towards Short Movie Generation,
in **ICLR** 2026.
Junfei Xiao, Ceyuan Yang, Lvmin Zhang, **Shengqu Cai**, Yang Zhao, Yuwei Guo,
Gordon Wetzstein, Maneesh Agrawala, Alan Yuille, Lu Jiang

[10] CL-Splats: Continual Learning of Gaussian Splatting with Local Optimization,
in **ICCV** 2025.
Jan Ackermann, Jonas Kulhanek, **Shengqu Cai**, Haofei Xu,
Marc Pollefeys, Gordon Wetzstein, Leonidas Guibas, Songyou Peng

[9] ByteMorph: Benchmarking Instruction-Guided Image Editing with Non-Rigid Motions,
in **arXiv** 2025.
Di Chang*, Mingdeng Cao*, Yichun Shi, Bo Liu, **Shengqu Cai**, Shijie Zhou,
Weilin Huang, Gordon Wetzstein, Mohammad Soleymani, Peng Wang.

[8] Diffusion Self-Distillation for Zero-Shot Customized Image Generation,
in **CVPR** 2025.
Shengqu Cai, Eric Ryan Chan, Yunzhi Zhang, Leonidas Guibas,
Jiajun Wu, Gordon Wetzstein.

[7] ReStyle3D: Scene-level Appearance Transfer with Semantic Correspondences,
in **SIGGRAPH** 2025.
Liyuan Zhu, **Shengqu Cai***, Shengyu Huang*, Gordon Wetzstein,
Naji Khosravan, Iro Armeni.

[6] X-Dyna: Expressive Dynamic Human Image Animation,
in **CVPR** 2025 (Highlight).
Di Chang, Hongyi Xu*, You Xie*, Yipeng Gao*, Zhengfei Kuang*, **Shengqu Cai***,
Chenxu Zhang*, Guoxian Song, Chao Wang, Yichun Shi, Zeyuan Chen,
Shijie Zhou, Linjie Luo, Gordon Wetzstein, Mohammad Soleymani.

[5] Collaborative Video Diffusion: Consistent Multi-video Generation with Camera Control,
in **NeurIPS** 2024.
Zhengfei Kuang*, **Shengqu Cai***, Hao He, Yinghao Xu, Hongsheng Li, Leonidas Guibas, Gordon Wetzstein.

[4] Robust Symmetry Detection via Riemannian Langevin Dynamics,
in **SIGGRAPH Asia** 2024.
Jihyeon Je*, Jiayi Liu*, Guandao Yang*, Boyang Deng*, **Shengqu Cai**, Gordon Wetzstein, Or Litany, Leonidas Guibas.

[3] Generative Rendering: Controllable 4D-Guided Video Generation with 2D Diffusion Models,
in **CVPR** 2024.
Shengqu Cai, Duygu Ceylan*, Matheus Gadelha*, Chun-Hao Paul Huang, Tuanfeng Y. Wang, Gordon Wetzstein.

[2] DiffDreamer: Towards Consistent Unsupervised Single-view Scene Extrapolation with Conditional Diffusion Models,
in **ICCV** 2023.
Shengqu Cai, Eric Ryan Chan, Songyou Peng, Mohamad Shahbazi, Anton Obukhov, Luc Van Gool, Gordon Wetzstein.

[1] Pix2NeRF: Unsupervised Conditional π-GAN for Single Image to Neural Radiance Fields Translation,
in **CVPR** 2022.
Shengqu Cai, Anton Obukhov, Dengxin Dai, Luc Van Gool.

Patent

[B] Diffusion-based Novel View Synthesis and Animation
US patent 18738823, approved in 2025.

[A] System for Unsupervised Single Image to Neural Radiance Fields Translation
European patent 4235588A1, approved in 2022.

Teaching Experience

2026	Deep Reinforcement Learning, Stanford University
2025	Continuous Mathematical Methods, Stanford University
2019	Practical Experiences Of Programming, King's College London

Academic services

Conference review:	CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML, SIGGRAPH, SIGGRAPH Asia
Journal review:	IJCV, Computing Surveys, Eurographics, T-PAMI