

Yuhe Zhong

Mobile: 07344303965 | yz2002@cam.ac.uk | www.linkedin.com/in/yuhe-zhong | github.com/Rainiver |

EDUCATION

University of Cambridge <i>MPhil in Data Intensive Science</i>	Cambridge, UK <i>Sep. 2025 – Aug. 2026 (Expected)</i>
Beihang University <i>Doctoral Student in Computer Science (quit)</i> <ul style="list-style-type: none">• Focus: Applied Machine Learning, Computer Vision, Generative AI	Beijing, China <i>Sep. 2022 – Jun. 2025</i>
Beihang University <i>B.Eng (Hons) in Electronic Engineering</i> <ul style="list-style-type: none">• GPA: 90.06/100 — Outstanding Graduate (Top 5%)	Beijing, China <i>Sep. 2018 – Jun. 2022</i>

SKILLS

Programming: Python, C/C++, MATLAB, Verilog, L^AT_EX
Developer Tools: Linux, Git, Docker, VS Code, PyCharm
Data/ML: PyTorch, SQL, Pandas, NumPy, TensorFlow, Scikit-learn
Web/Full-Stack: React, Node.js, MongoDB, Express

WORK EXPERIENCE

Alibaba Group <i>Visual Generation Research Intern</i> <ul style="list-style-type: none">• Developed large-scale multi-view human image synthesis datasets (1000 identities, 50 poses, 5 viewpoints) using fine-tuned diffusion models• Built a framework for 3D avatar generation from a single image, combining diffusion-based multi-view synthesis with transformer-based SDF reconstruction and refining geometry using normal maps• Open-sourced preprocessing and reconstruction toolkits: (AvatarMesh, PrePose, Human Datasets Preprocessor,)	Beijing, China <i>Oct. 2023 – Dec. 2024</i>
SenseTime <i>AI Video Codec Research Intern</i> <ul style="list-style-type: none">• Implemented deep learning-based variational image compression models with PyTorch, integrating multiple state-of-the-art methods into a modular and reusable framework• Optimized ROI-based compression, achieving 2% PSNR and 1.8% MS-SSIM improvement over baseline models• Open-sourced reorganized version of this work for reproducibility: Deep Learning Image Compression	Beijing, China <i>Oct. 2021 – Mar. 2022</i>

RESEARCH & DEVELOPMENT EXPERIENCE

Text-to-3D Complex Scene Generation using Gaussian Splatting <ul style="list-style-type: none">• Built a system for generating 3D scenes from text prompts by combining Gaussian Splatting and LLMs for semantic guidance• Implemented local-global training strategies, progressive scale control, and collision loss to ensure scene consistency and scalability• Achieved 4% higher CLIP similarity compared to SOTA, demonstrating improved semantic fidelity	<i>Mar. 2024 – Jul. 2024</i>
3D Human Body Reconstruction and Animation via Diffusion Models <i>Core Member, Beijing Natural Science Foundation Program</i> <ul style="list-style-type: none">• Developed end-to-end pipelines for 3D human body reconstruction and animation using 3D-aware diffusion models with SMPL priors• Designed two-stage denoising process combining 3D-aware and 2D image denoisers for better generalization to unseen poses• Improved FID by 3% and LPIPS by 7% on ZJU-MoCap dataset compared to prior methods	<i>Mar. 2023 – Jan. 2024</i>
Face Video Synthesis with Neural Radiance Fields <i>Bachelor's Thesis (English simplified version[pdf])</i> <ul style="list-style-type: none">• Developed full pipeline for face video synthesis with Neural Radiance Fields (NeRF), integrating 3D Morphable Models and semantic parsing for geometric priors• Implemented separate NeRFs for head and torso, embedding facial identity to generalize across identities• Achieved 5% PSNR improvement over SOTA face reenactment methods	<i>Dec. 2021 – May 2022</i>