Yuhe Zhong

+44 7344 303965 | yz2002@cam.ac.uk | www.linkedin.com/in/yuhe-zhong | rainiver.github.io |

EDUCATION

University of Cambridge — Cambridge, UK
MPhil in Data Intensive Science — *Sep 2025 – Aug 2026 (Expected)*

Beihang University — Beijing, China
Doctoral Student in Computer Science (withdrew) — *Sep 2022 – Jun 2025*

- **Focus**: Applied Machine Learning, Computer Vision, Generative AI

Beihang University — Beijing, China
B.Eng (Hons) in Electronic Engineering — *Sep 2018 – Jun 2022*

- **GPA**: 90.06/100 — Outstanding Graduate (Top 5%)

SKILLS

Languages: Python, C/C++, JavaScript, SQL
ML/AI: PyTorch, TensorFlow, Scikit-learn, Hugging Face, LangChain, RAG
Data/Backend: MongoDB, Supabase, Prisma ORM, Docker
Web/DevOps: Next.js, React, Node.js, Vercel CI/CD, Git

WORK EXPERIENCE

Alibaba Group | Visual Generation Research Intern — *Oct 2023 – Dec 2024*

- Developed large-scale multi-view human image synthesis datasets (1000 IDs, 50 poses, 5 views) 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)

SenseTime | AI Video Codec Research Intern — *Oct 2021 – Mar 2022*

- 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

PROJECTS

AI-Powered Travel Booking Platform [code] [demo] — *Sep 2025 – Oct 2025*

- Developed full-stack platform (*Next.js, TypeScript, Prisma, MongoDB*) integrating **RAG pipeline** and **multi-agent LLM system**
- Implemented **vector semantic search** (768-dim embeddings, Supabase pgvector) using **Gemini Pro API**
- Built **dynamic pricing model** and **context-aware reasoning memory** for personalized booking
- Deployed production-ready CI/CD on **Vercel**, achieving 1–3s latency and high retrieval precision

Text-to-3D Complex Scene Generation using Gaussian Splatting — *Mar 2024 – Jul 2024*

- 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 for scene consistency
- Achieved 4% higher CLIP similarity compared to SOTA, demonstrating improved semantic fidelity

3D Human Body Reconstruction and Animation via Diffusion Models — *Mar 2023 – Jan 2024*

- 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 the FID by 3% and LPIPS by 7% on the ZJU-MoCap dataset over previous methods

Face Video Synthesis with Neural Radiance Fields [pdf] — *Dec 2021 – May 2022*

- Developed full pipeline for face video synthesis with **Neural Radience Fields** (NeRF), integrating **3D Morphable Models** and semantic parsing for geometric priors
- Implemented separate NeRFs for head and torso, embedding facial identity vectors to enhance cross-identity generalization
- Achieved 5% PSNR improvement over SOTA face reenactment methods