

Hanzhang Tu

PhD Candidate | 3D Vision, Neural Rendering & Real-Time 3D Systems

Beijing, China · thz22@mails.tsinghua.edu.cn · +86 138 2998 2493

[Google Scholar](#)

Research profile

Ph.D. candidate at Tsinghua University specializing in 3D/4D vision and neural rendering. Builds end-to-end research systems spanning data synthesis, model training, GPU optimization, and real-time deployment. Seeking a Research or Applied Scientist role in **world models** and **3D/4D vision**.

Education & research appointment

PhD in Control Science and Engineering

Aug 2022 - Jun 2027

Department of Automation, Tsinghua University

Beijing, China

Expected completion: June 2027. Advisor: [Prof. Yebin Liu](#).

BEng in Measurement and Control Technology and Instrumentation

Aug 2018 - Jul 2022

Department of Precision Instrument, Tsinghua University

Beijing, China

Minor in Computer Science.

Selected research projects

Tele360: Real-Time Feed-Forward Human Reconstruction from Sparse Unposed Cameras

TOG / SIGGRAPH Asia 2026 · Dynamic human reconstruction, freeview synthesis, real-time streaming

- Built a feed-forward pipeline that jointly estimates camera poses and reconstructs dynamic 3D Gaussians from 4–6 sparse, unposed 2K RGB streams.
- Delivered input-to-rendering at over 25 FPS on one consumer GPU, with Gaussian maps streamed to remote devices through standard video codecs.
- Iterative free-view rendering on various portable devices.

GBC-Splat: Generalizable Gaussian-Based Clothed Human Digitalization under Sparse RGB Cameras

CVPR 2025 · Clothed-human digitalization, 3D Gaussians, sparse-view reconstruction

- Reconstructed complete clothed-human geometry and Gaussian appearance directly from sparse RGB input without subject-specific optimization.
- Combined implicit occupancy and cross-view disparity estimation to support loose clothing and photorealistic 6-DoF novel-view synthesis without an SMPL template.

Tele-Aloha: A Low-budget and High-authenticity Telepresence System Using Sparse RGB Cameras

SIGGRAPH 2024 · Real-time telepresence, free-view synthesis, human performance rendering

- Designed an end-to-end capture, reconstruction, streaming, and autostereoscopic display pipeline using four cameras and one consumer GPU.
- Achieved 2048×2048 rendering at 30 FPS with end-to-end latency below 150 ms for bidirectional high-authenticity remote communication.

Internship experience

Tencent

Jul 2024 - Aug 2025

Large-model inference systems

- Implemented KV cache CPU offloading for mixed long- and short-generation workloads, improving large-model inference throughput on production data by 10%–230%.
- Extended the offloading path to tensor and pipeline parallelism for multi-GPU inference, broadening compatibility across distributed deployment configurations.
- Deployed the feature into the production codebase after rigorous cross-hardware validation.

Technical Skills

3D Vision	3D/4D Reconstruction, Neural Rendering (NeRF, 3DGS), Multi-view Geometry, Dynamic Human/Scene Modeling, Geometric Foundation Model & World Models
Machine Learning	PyTorch, Triton, Multi-GPU Distributed Training, Synthetic Data Generation (Blender, 3ds Max), Feed-forward Architectures
Systems Deployment	GPU Optimization (Custom CUDA Kernels), Real-time Rendering (OpenGL, Metal), Inference Acceleration (TensorRT, ONNX), Low-latency Streaming, Multi-sensor Synchronization & Calibration
Programming & Tools	Python, C++, CUDA, Rust, Docker, Git

Selected publications

Hanzhang Tu, Z. Liao, W. Min, J. Zhang, Y. Liu. **Tele360: Real-Time Feed-Forward Human Reconstruction from Sparse Unposed Cameras**. *ACM Transactions on Graphics (SIGGRAPH Asia 2026 Journal Track)*, accepted.

Jointly estimates camera poses and predicts dynamic 3D Gaussians from sparse unposed streams, enabling real-time 2K input-to-rendering on one consumer GPU.

Z. Liao, Hanzhang Tu, et al. **HADES: Human Avatar with Dynamic Explicit Hair Strands**. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2025.

Represents hair as strands bound to 3D Gaussians and models inertial, velocity-aware motion.

Hanzhang Tu, Z. Liao, et al. **GBC-Splat: Generalizable Gaussian-Based Clothed Human Digitalization under Sparse RGB Cameras**. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

Reconstructs complete human shapes and Gaussian splats from sparse-view RGB inputs in a feed-forward manner, including loose garments.

B. Zhou, S. Zheng, Hanzhang Tu, et al. **GPS-Gaussian+: Generalizable Pixel-wise 3D Gaussian Splatting for Real-Time Human-Scene Rendering from Sparse Views**. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

Directly regresses source-view Gaussian parameter maps for instant sparse-view novel-view synthesis without per-subject fine-tuning or optimization.

Hanzhang Tu, R. Shao, et al. **Tele-Aloha: A Low-budget and High-authenticity Telepresence System Using Sparse RGB Cameras**. *ACM SIGGRAPH*, 2024.

Combines four RGB inputs, cascaded disparity estimation, 3D Gaussian neural rasterization, and weighted blending for real-time 2K telepresence.

R. Shao, Z. Zheng, Hanzhang Tu, et al. **Tensor4D: Efficient Neural 4D Decomposition for High-Fidelity Dynamic Reconstruction and Rendering**. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR Highlight)*, 2023.

Hierarchically decomposes 4D spatiotemporal tensors into time-aware volumes and compact feature planes for memory-efficient dynamic scene modeling.

Full publication list: [Google Scholar](#).

Academic service

Reviewer for CVPR, ICCV, ECCV, NeurIPS, IEEE Transactions on Instrumentation and Measurement (TIM), and IEEE Transactions on Circuits and Systems for Video Technology (TCSVT).

Languages

Mandarin Chinese (native); English (CET-4: 625, CET-6: 594); Japanese (JLPT N2).