Zhongyuan Hu

Phone: +86 153-2058-1968 | Email: huzhongyyuan@gmail.com
Website: huzhongyyuan.github.io | Google Scholar: HBE737kAAAAJ
Hometown: Chongqing, China | Research: Generative AI, 3D Human Motion, Video Generation, Multimodal Learning

EDUCATION

Tsinghua University **Ph.D. in Control Science and Engineering** Sep 2025 - Jun 2030

- Direct-entry Ph.D. student; currently in the first year of study

Chongqing University **B.Eng. in Computer Science and Technology** Sep 2021 - Jun 2025

- **GPA: 3.91/4.00; ranked 1/323 (top 0.3%)**
- **Honors:** National Scholarship (top 1%); First-Class Scholarship (top 3%)

RESEARCH EXPERIENCE

1. InfiniteDance: Scalable 3D Dance Generation Towards in-the-wild Generalization (Co-first Author, Core Contributor, ECCV 2026) Mar 2025 - Aug 2025

Keywords: Motion Generation; LLM; Human Animation; Motion Data Pipeline

Project: Addressed weak generalization to unseen music, disordered motion semantics, and foot-related physical artifacts. Made core contributions to an automated monocular motion-reconstruction pipeline and a foot-restoration diffusion model, producing over 100 hours of multimodal 3D dance data. Co-developed ChoreoLLaMA with RAG for unseen music and a slow-fast Mixture-of-Experts module for tempo-adaptive generation.

Paper: arxiv.org/abs/2603.13375 | Project: infinitedance.github.io

2. Camera Operator: Object-Grounded Camera Trajectory Generation from Text and 3D Bounding Box Sequences (First Author, Core Contributor, ACM MM 2026) Jan 2026 - Apr 2026

Keywords: Camera Generation; Flow Matching; Video Generation

Project: Built BlockCam, a 41K-sequence benchmark combining Unreal Engine and real-video pipelines with aligned text, 3D bounding-box state sequences, and camera trajectories. Proposed a geometry-conditioned flow-matching model with look-at, framing, visibility, and temporal supervision to mitigate composition drift and maintain stable long-horizon camera-subject interaction.

Paper: doi.org/10.1145/3767308.3835459

3. Controllable Video Generation: A Survey (Co-first Author, Core Contributor, under review at ACM Computing Surveys) Apr 2025 - Aug 2026

Keywords: Controllable Video Generation; Diffusion Models; Multimodal Conditioning

Project: Analyzed how camera motion, depth, human pose, and other non-text conditions enter the denoising process of video diffusion models. Organized the literature into single-condition, multi-condition, and universal controllable generation, and co-maintained the Awesome-Controllable-Video-Generation repository.

Paper: arxiv.org/abs/2507.16869

INTERNSHIP EXPERIENCE

Tencent Games **Research Intern, Shenzhen** Jun 2025 - Dec 2025

WATCH: World-aware Allied Trajectory and pose reconstruction for Camera and Human (Co-first Author, Core Contributor, Accepted)

Keywords: Video Motion Capture; Human Pose Estimation; Camera Estimation

Project: Addressed strongly coupled human-camera motion and underused camera rotation and translation cues in in-the-wild monocular reconstruction. Contributed heading-angle decomposition and global-velocity modeling to reduce long-sequence drift. Outperformed GVHMR and WHAM; replacing GVHMR's corresponding module further improved dynamic-camera reconstruction. The algorithm was deployed in Tencent products.

Paper: arxiv.org/abs/2509.04600