Gong Kehong

Singapore (+65) 8432 1426 gongkehong@u.nus.edu
garfield-kh.github.io github.com/Garfield-kh Google Scholar

Summary

AI researcher in **generative AI, visual generation, and efficient deep learning**. Five first-author papers at top venues (CVPR×3, ICCV, TPAMI), including **2 Orals** and a **CVPR Best Paper Candidate**. My work spans video and world models, 3D motion generation and capture, **model-level LLM/VLM inference acceleration**, and deployment-oriented AI research, including on-device motion capture and an on-device LLM released with a carrier partner.

Research Interests

Generative and multimodal foundation models; video generation and world models; 3D human and animal motion generation and capture; efficient deep learning (low-bit quantization, sparse/efficient attention, model compression); and digital humans and embodied/physical AI.

Technical Skills

Programming: Python. **Frameworks:** PyTorch; diffusion- and transformer-based generative modeling.
Model efficiency: low-bit quantization, sparse/efficient attention, model compression, and model-level inference acceleration.
Research areas: visual generation, video/world models, 3D motion generation and capture, LLM/VLM foundation models, digital humans, physical AI.

Education

Jan 2019 – Dec 2022 **Ph.D. in Electrical and Computer Engineering**, National University of Singapore

- Dissertation: *Deep Learning in Human Pose Generation and Its Application.*
- Supervisors: Asst. Prof. Jiashi Feng (Year 1–3) and Assoc. Prof. Xinchao Wang (Year 3–4).

Aug 2013 – May 2017 **B.Eng. (Honours), Engineering Science Programme**, National University of Singapore

- Specialisation: Computational Engineering Science.

Professional Experience

Dec 2022 – Present **Senior Staff Engineer**, Huawei International Pte. Ltd. (Singapore)

- **2026 – present:** Pre-research on generative video models, including world models and joint audio–video generation.
- **2025:** Built unified 3D motion capture & retargeting for arbitrary skeletons from monocular videos, toward scalable motion foundation models for digital humans and embodied agents; *deployed in the virtual-pet feature of Huawei smartphones.* Published as the *MoCapAnything* series (CVPR 2026).
- **2023–2024: Model-level efficiency and deployment of LLMs/VLMs**: low-bit quantization, sparse/efficient attention, and model compression. Our on-device LLM for smart-home terminals was *released at the China Mobile Global Partners Conference 2023.*

- **Mentoring:** Supervised research interns; mentored work published as *MotionMix* (AAAI 2024, as corresponding author).

Publications

Summary: 5 first-author + 1 corresponding-author papers published at top-tier venues (CVPR×3, ICCV, TPAMI, AAAI), with 2 additional first-author papers under review. Highlights: **2 Orals** (CVPR 2021, CVPR 2022) and **1 CVPR Best Paper Candidate** (CVPR 2021).

First-Author Publications

- **MoCapAnything: Unified 3D Motion Capture for Arbitrary Skeletons from Monocular Videos**
  Gong Kehong, Zhengyu Wen, Weixia He, Mingxi Xu, Qi Wang, Ning Zhang, Zhengyu Li, Dongze Lian, Wei Zhao, Xiaoyu He, Mingyuan Zhang
  CVPR 2026 – arXiv:2512.10881
- **MoCapAnything V2: End-to-End Motion Capture for Arbitrary Skeletons**
  Gong Kehong, Zhengyu Wen, Dao Thien Phong, Mingxi Xu, Weixia He, Qi Wang, Ning Zhang, Zhengyu Li, Guanli Hou, Dongze Lian, Xiaoyu He, Mingyuan Zhang, Hanwang Zhang
  Under Review, 2026 – arXiv:2604.28130
- **SWiT-4D: Sliding-Window Transformer for Lossless and Parameter-Free Temporal 4D Generation**
  Gong Kehong, Zhengyu Wen, Mingxi Xu, Weixia He, Qi Wang, Ning Zhang, Zhengyu Li, Chenbin Li, Dongze Lian, Wei Zhao, Xiaoyu He, Mingyuan Zhang
  Under Review, 2025 – arXiv:2512.10860
- **TM2D: Bimodality Driven 3D Dance Generation via Music-Text Integration**
  Gong Kehong, Dongze Lian, Heng Chang, Chuan Guo, Zihang Jiang, Xinxin Zuo, Michael Bi Mi, Xinchao Wang
  ICCV 2023 – arXiv:2304.02419
- **Learning to Augment Poses for 3D Human Pose Estimation in Images and Videos**
  Jianfeng Zhang*, **Gong Kehong***, Xinchao Wang, Jiashi Feng (*equal contribution)
  IEEE TPAMI, 2023 (co-first author) – IEEE Xplore
- **PoseTriplet: Co-evolving 3D Human Pose Estimation, Imitation, and Hallucination under Self-supervision**
  Gong Kehong, Bingbing Li, Jianfeng Zhang, Tao Wang, Jing Huang, Michael Bi Mi, Jiashi Feng, Xinchao Wang
  CVPR 2022 (Oral) – arXiv:2203.15625
- **PoseAug: A Differentiable Pose Augmentation Framework for 3D Human Pose Estimation**
  Gong Kehong, Jianfeng Zhang, Jiashi Feng
  CVPR 2021 (Oral, Best Paper Candidate) – arXiv:2105.02465

Corresponding-Author Publications

- **MotionMix: Weakly-Supervised Diffusion for Controllable Motion Generation**
  Nhat M. Hoang, **Gong Kehong** (corresponding author), Chuan Guo, Michael Bi Mi
  AAAI 2024 – arXiv:2401.11115

Co-Authored Publications

- **DiMo: Discrete Diffusion Modeling for Motion Generation and Understanding**
  Ning Zhang, Zhengyu Li, Kwong Weng Loh, Mingxi Xu, Qi Wang, Zhengyu Wen, Xiaoyu He, Wei Zhao, **Gong Kehong**, Mingyuan Zhang
  arXiv preprint, 2026 – arXiv:2602.04188

- **NECromancer: Breathing Life into Skeletons via BVH Animation**
  Mingxi Xu, Qi Wang, Zhengyu Wen, Phong Dao Thien, Zhengyu Li, Ning Zhang, Xiaoyu He, Wei Zhao, **Gong Kehong**, Mingyuan Zhang
  arXiv preprint, 2026 – arXiv:2602.06548
- **Priority-centric Human Motion Generation in Discrete Latent Space**
  Hanyang Kong, **Gong Kehong**, Dongze Lian, Michael Bi Mi, Xinchao Wang
  ICCV 2023 – arXiv:2308.14480

Honors & Awards

- **Best Paper Candidate**, CVPR 2021 (PoseAug).
- Oral Presentation, CVPR 2021 & CVPR 2022.

Academic Service

Reviewer: CVPR, ECCV, NeurIPS, AAAI, ACM MM, Pattern Recognition, Robotics and Autonomous Systems.