

TAEWOONG KANG

Phone: (+082) 10-5481-4597 ◇ Email: keh0t0@kaist.ac.kr

Homepage: keh0t0.github.io

Google Scholar ◇ Github

EDUCATION

Korea Advanced Institute of Science and Technology (KAIST)

March 2024 - Present

Integrated M.S./Ph.D. program in Artificial Intelligence

Expected Feb. 2029

Current GPA: 3.86/4.3

Advisor : Jaegul Choo

Korea University

March 2018 - Feb 2024

B.E. in Electoronical Engineering

GPA: 4.01/4.5

Major GPA: 4.24/4.5

RESEARCH INTERESTS

My research aims to build general-purpose robots capable of dexterous manipulation in the physical world. I study how large-scale generative models, such as video world models and world-action models, can give robots the physical understanding, memory, and adaptability needed to perform diverse, contact-rich tasks.

Keywords : Robot Learning, Dexterous Manipulation, World Action Model, Video World Model, Embodied AI

ONGOING PROJECTS

In-Context Self-Improvement for World-Action Models. Robots deployed in new environments often fail on their first attempts, and adapting them typically requires costly fine-tuning. We train a large-scale world-action model to improve from its own trial and error in context: conditioned on a few recent attempts and their outcomes, the policy corrects its next actions without gradient updates, and can additionally take a human demonstration video to learn new tasks in one shot.

Real-to-Sim Task Generation for Dexterous Manipulation. Training general-purpose dexterous hands with reinforcement learning requires a large and diverse set of simulated tasks. As part of a mega-scale dexterous RL project, we convert real-world scenes and objects into simulation to generate diverse manipulation tasks at scale.

Scene-Level Long-Term Memory Benchmark for Robot Manipulation. Long-horizon household tasks require robots to recall scene-level information observed far in the past, yet existing benchmarks rarely isolate this capability. We introduce a benchmark spanning diverse situations and memory types to rigorously measure long-term memory in robot policies.

Multi-View Consistent Video Generation for World-Action Models. World-action models increasingly rely on multi-view video generation for policy learning, but generated views often lack geometric consistency. We enforce cross-view consistency in generated videos to improve robot policy performance.

PUBLICATIONS

EgoX: Egocentric Video Generation from a Single Exocentric Video. [1]

Taewoong Kang*, Kinam Kim*, Dohyeon Kim, Minho Park, Junha Hyung, Jaegul Choo. CVPR, 2026, Denver, USA.

AHS: Adaptive Head Synthesis via Synthetic Data Augmentation.

Taewoong Kang*, Hyojin Jang*, Sohyun Jeong*, Jaeseong Lee, Seunggi Moon, Gihwi Kim, Hoon Jin Jung, Jaegul Choo. CVPR, 2026, Denver, USA.

SphereDiff: Tuning-free Omnidirectional Panoramic Image and Video Generation via Sphere Latent Representation.[2]

Minho Park*, **Taewoong Kang***, Jooyeol Yun, Sungwon Hwang, Jaegul Choo. Association for the Advancement of Artificial Intelligence (AAAI), 2026, Singapore. (Acceptance Rate: 17.6%, oral presentation)

TexAvatars: Hybrid Texel-3D Representations for Stable Rigging of Photorealistic Gaussian Head Avatars.[3]

Jaeseong Lee*, Junyeung Ahn*, **Taewoong Kang**, Jaegul Choo. International Conference on 3D Vision (3DV), 2026 Vancouver, Canada.

SurFhead: Affine Rig Blending for Geometrically Accurate 2D Gaussian Surfel-based Head Avatars [4]

Jaeseong Lee*, **Taewoong Kang***, Marcel Bühler, Minjung Kim, Sungwon Hwang, Junha Hyung, Hyojin Jang and Jaegul Choo. ICLR, 2025, Singapore.

Zero-Shot Head Swapping in Real-World Scenarios.[5]

Taewoong Kang*, Sohyun Jeong*, Hyojin Jang*, Jaegul Choo. CVPR, 2025, Nashevill, USA.

VEGS: View Extrapolation of Urban Scenes in 3D Gaussian Splatting using Learned Priors

Sungwon Hwang*, Minjung Kim*, **Taewoong Kang**, Jayeon Kang, and Jaegul Choo. European Conference on Computer Vision (ECCV), 2024, Milano, Italy.

Expression Domain Translation Network for Cross-domain Head Reenactment [6]

Taewoong Kang*, Jeongsik Oh*, Jaeseong Lee, Sunghyun Park, and Jaegul Choo International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024, Seoul, Korea.

Injecting Inductive Bias to 3D Gaussian Splatting for Geometrically Accurate Radiance Fields

Sungwon Hwang, Soohyun Ryu, Jaeseong Yoon, Jaeseong Lee, Junha Hyung, Min-Jung Kim, **Taewoong Kang**, Joohyun Park, Jaegul Choo, Suyoung Yeon.

GauShot: One-shot Gaussian Head Avatars via Expressive Diffusion Reenactment Prior

Jaeseong Lee, **Taewoong Kang**, Nikita Drobyshev, Junho Kim, Jaegul Choo, Gayoung Lee, Jin-Hwa Kim.

ACHIEVEMENTS

Sukrim Scholarship Korea Uni.

Fall 2022

An award certificate

Fall 2023, Spring 2023, Fall 2022, Spring 2022

SKILLS

Programming Languages

Python, C/C++, MATLAB

Machine Learning Tools

Pytorch, Sklearn, Numpy

- [1] **Taewoong Kang**, K. Kim, D. Kim, M. Park, J. Hyung, and J. Choo, *Egox: Egocentric video generation from a single exocentric video*, 2025. arXiv: 2512.08269 [cs.CV]. [Online]. Available: <https://arxiv.org/abs/2512.08269>.
- [2] M. Park, **Kang, Taewoong**, J. Yun, S. Hwang, and J. Choo, “Spherediff: Tuning-free omnidirectional panoramic image and video generation via spherical latent representation,” *arXiv preprint arXiv:2504.14396*, 2025.
- [3] J. Lee, J. Ahn, **Kang, Taewoong**, and J. Choo, “Texavatars: Hybrid texel-3d representations for stable rigging of photorealistic gaussian head avatars,” *arXiv preprint arXiv:2512.21099*, 2025.
- [4] J. Lee, **Taewoong Kang**, M. Buehler, *et al.*, “Surfhead: Affine rig blending for geometrically accurate 2d gaussian surfel head avatars,” in *The Thirteenth International Conference on Learning Representations*, 2025.

- [5] **Kang, Taewoong**, S. Jeong, H. Jang, and J. Choo, “Zero-shot head swapping in real-world scenarios,” in *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2025, pp. 10 805–10 814.
- [6] **Kang, Taewoong**, J. Oh, J. Lee, S. Park, and J. Choo, “Expression domain translation network for cross-domain head reenactment,” in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2024, pp. 7356–7359.