Ho Kei (Rex) Cheng

Email: hkchengrex [at] gmail.com, or, hokeikc2 [at] illinois.edu

hkchengrex.com

GitHub: github.com/hkchengrex

Education

University of Illinois Urbana-Champaign, Illinois, United States — 2026
PhD in Computer Science
Advisor: Alexander Schwing

The Hong Kong University of Science and Technology, Hong Kong — 2021
MPhil in Computer Science
Advisors: Chi-Keung Tang and Yu-Wing Tai

The Hong Kong University of Science and Technology, Hong Kong — 2019
BEng in Computer Science, with a minor in Robotics

Selected Publications

SAM 3: Segment Anything with Concepts — ICLR '26
Nicolas Carion, Laura Gustafson, Yuan-Ting Hu, ..., **Ho Kei Cheng**, et al.

ImagenWorld: Stress-Testing Image Generation Models with Explainable Human Evaluation on Open-ended Real-World Tasks — ICLR '26
Samin Mahdizadeh Sani, Max Ku, Nima Jamali, ..., **Ho Kei Cheng**, Ping Nie, Wenhu Chen.

The Curse of Conditions: Analyzing and Improving Optimal Transport for Conditional Flow-Based Generation — ICCV '25
Ho Kei Cheng, Alexander Schwing.

MMAudio: Taming Multimodal Joint Training for High-Quality Video-to-Audio Synthesis — CVPR '25
Ho Kei Cheng, Masato Ishii, Akio Hayakawa, Takashi Shibuya, Alexander Schwing, Yuki Mitsufuji. *Media*.

Putting the Object Back into Video Object Segmentation — CVPR '24
Ho Kei Cheng, Seoung Wug Oh, Brian Price, Joon-Young Lee, Alexander Schwing. *Used by iMotions and Annolid.*

Tracking Anything with Decoupled Video Segmentation — ICCV '23
Ho Kei Cheng, Seoung Wug Oh, Brian Price, Alexander Schwing, Joon-Young Lee.

XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model — ECCV '22
Ho Kei Cheng, Alexander Schwing. *Used by supervisely, Track-Anything, and more.*

Rethinking Space-Time Networks with Improved Memory Coverage for Efficient Video Object Segmentation — NeurIPS '21
Ho Kei Cheng, Yu-Wing Tai, Chi Keung Tang. *Used by Trioscope and BURST.*

Modular Interactive Video Object Segmentation: Interaction-to-Mask, Propagation and Difference-Aware Fusion — CVPR '21
Ho Kei Cheng, Yu-Wing Tai, Chi Keung Tang. *Used by Sieve.*

CascadePSP: Toward Class-Agnostic and Very High-Resolution Segmentation via Global and Local Refinement — CVPR '20
Ho Kei Cheng*, Jihoon Chung*, Yu-Wing Tai, Chi Keung Tang.

Work Experience

Research Scientist, Luma AI, Redwood City, United States — May 2026 - Present
Working on multimodal foundation models.

Research Intern, Meta FAIR/MSL, San Francisco, United States — May 2025 - Nov 2025
Developed object multiplexing for efficient video inference in SAM 3 and SAM 3.1.

Research Intern, Sony AI, Tokyo, Japan — May 2024 - Nov 2024
Developed MMAudio, a multimodal flow-matching generative model; trained from scratch.

Research Intern, Kaiber AI, Remote, United States — Nov 2023 - May 2024
Researched controllable video generation and editing pipelines using diffusion models.

Research Intern, Adobe Research, Remote, United States — May 2022 - Nov 2022
Developed DEVA and Cutie for efficient video segmentation; technology contributed to Adobe Express and After Effects.

Invited Talks

Object-Level Reasoning in Video Object Segmentation and Its Multimodal Applications, Twelve Labs 2024

Segmenting Videos in the Open World, IBM Zurich, Accelerated Discovery 2023

Large-Scale Decoupled Video Segmentation, Apple 2023

Professional Activities

- Co-organized the 1st Workshop on Generative AI for Audio-Visual Content Creation (Gen4AVC) at ICCV 2025.
- Conference Reviewer: CVPR, ICCV, ECCV, NeurIPS, ICLR, ICML, AAAI, SIGGRAPH Asia.
- Journal Reviewer: TMLR, IEEE TPAMI, IEEE TIP, IEEE TCSVT, IEEE/ACM TASLPRO, Pattern Recognition.
- Outstanding Reviewer: ICML 2022, CVPR 2026.