Gene Chou

gene@cs.cornell.edu · genechou.com · github.com/gene-chou

EDUCATION

Cornell University 2023 -
PhD student in computer science
Advisors: Noah Snavely and Bharath Hariharan

Princeton University 2018 - 2022
BSE in computer science, minor in applied mathematics
Honors: magna cum laude, departmental GPA: 3.96 / 4.0

RESEARCH

Adobe Research | *San Jose, CA* 5/2024 - 8/2024
Generated videos from unposed, wide-baseline images such as internet photos.

Backflip AI | *San Francisco, CA* 7/2023 - 9/2023
Scaled up 3D asset generation at a GenAI startup.

Disney Research Imagineering | *Glendale, CA* 2/2023 - 7/2023
Combined large vision models and RL for human-robot interaction and sim-to-real transfer.

Princeton Computational Imaging Lab | *Princeton, NJ* 3/2022 - 2/2023
Explored generalization and generation of neural scene representations, advised by Prof. Felix Heide.

Princeton Visual AI Lab | *Princeton, NJ* 6/2021 - 12/2021
Stress-tested algorithmic fairness via synthetic data generation, advised by Prof. Olga Russakovsky.

Multimedia Technology Lab at Academia Sinica | *Taipei, Taiwan* 6/2020 - 8/2021
Improved object detection robustness and efficiency, advised by Prof. Hong-Yuan Mark Liao.

EXPERIENCE

Teaching Assistant
Deep Learning (2024); Algorithms and Data Structures (2019 - 2022)

Service
Reviewer for CVPR, ICCV, ECCV, SIGGRAPH, TPAMI

Awards
NSF Graduate Research Fellowship (2023 - 2028)

PUBLICATIONS

Generating 3D-Consistent Videos from Unposed Internet Photos
G. Chou, K. Zhang, S. Bi, H. Tan, Z. Xu, F, Luan, B. Hariharan, N. Snavely *Arxiv 2024*
We propose the task of generating videos from sparse, unposed internet photos, and design a self-supervised method that takes advantage of the consistency of videos and variability of multiview internet photos to train a 3D-aware video model without any 3D annotations such as camera parameters.

MegaScenes: Scene-Level View Synthesis at Scale
J. Tung*, **G. Chou***, R. Cai, G. Yang, K. Zhang, G. Wetzstein, B. Hariharan, N. Snavely *ECCV 2024*
MegaScenes is a scene-level dataset containing 100K SfM reconstructions and 2M registered images, collected from Wikimedia Commons. We validate its effectiveness in training large-scale, generalizable models on the task of single image novel view synthesis.

Generalist YOLO: Towards Real-Time End-to-End Multi-Task Visual Language Models
HS. Chang, CY. Wang, R. Wang, **G. Chou**, HY. Liao *WACV 2025*
Builds on YOLOR to jointly train multiple vision and vision-language tasks. Fast and lightweight.

Thin On-Sensor Nanophotonic Array Cameras *SIGGRAPH Asia 2023*
P. Chakravarthula, J. Sun, X. Li, C. Lei, **G. Chou**, M. Bijelic, J. Froesch, A. Majumdar, F. Heide
Recovers images in broadband using a single flat metasurface optic and probabilistic deconvolution.

Diffusion-SDF: Conditional Generative Modeling of Signed Distance Functions
G. Chou, Y. Bahat, F. Heide *ICCV 2023*
Performs diffusion on the latent space of neural SDFs to conditionally generate 3D objects.

GenSDF: Two-Stage Learning of Generalizable Signed Distance Functions
G. Chou, I. Chugunov, F. Heide *NeurIPS 2022*
Reconstructs neural SDFs from raw input point clouds of over a hundred unseen object classes.