Chen Huang

Research Engineer, Apple

Generative AI: image, video, 3D | Multimodal LLM | On-device

San Jose, CA
(573) 823-8553
huangchen@outlook.com
chenhuang.me
linkedin.com/in/chuang5
Google Scholar

Research engineer at Apple with 9+ years of experience in computer vision and generative modeling, spanning image, video, and 3D generation, pretraining, fine-tuning, and efficient inference. Owned end-to-end training and distillation for Spatial Reframing, a shipped on-device Apple Photos feature, and co-led Cavia (ICML 2025).

Experience

Apple · Research Engineer

Cupertino, CA · Sep 2017 – Present

Spatial Reframing
On-device generative imaging, Apple Photos (iOS 27)

- Owned the end-to-end training lifecycle for the generative image model powering Spatial Reframing, shipped in Apple Photos for post-capture viewpoint adjustment.
- Designed the data collection pipeline and pretraining recipe; performed supervised fine-tuning to synthesize newly visible regions while preserving consistency with the original scene.
- Distilled the model for production on-device inference.

Video Generation
Cavia · ICML 2025
Camera-controlled multi-view diffusion

- Co-led research and experimentation for Cavia (ICML 2025), a diffusion framework generating multiple videos from a single image with explicit camera control and consistent scene dynamics.
- Developed view-integrated attention by extending spatial and temporal attention to improve cross-view and temporal consistency while preserving object motion.
- Jointly trained on static-scene, synthetic multi-view object, and real monocular videos; outperformed prior methods in geometric consistency and perceptual quality.

3D Generation
Image-conditioned 3D asset synthesis

- Designed the generation algorithm and model architecture for image-conditioned 3D asset synthesis, producing geometry and texture from a single reference image.
- Trained the generative model exclusively on internal synthetic 3D assets.

Face ID
Model compression and efficient inference

- Owned compression and deployment of Face ID recognition models on Apple hardware, reducing model size by **75%** while maintaining recognition accuracy.
- Optimized on-device inference through mixed-precision post-training quantization, low-bit weight representations, and operator fusion.

Object Capture
Photogrammetry and 3D reconstruction, RealityKit

- Developed core algorithms for the photogrammetry pipeline powering Apple Object Capture, shipped in RealityKit on iOS and macOS.
- Enabled reconstruction of textured 3D models from overlapping photographs for AR applications and spatial content creation.

Earlier Research Experience

Internships
Research

- **Microsoft Research · Computer Vision Research Intern**
  Redmond, WA · Mar – Sep 2017
  Developed a real-time pedestrian detection model and deployed it on custom Intel hardware for integration into Microsoft products.
- **Apple · Deep Learning Research Intern**
  Cupertino, CA · May – Sep 2016
  Distilled compact face recognition networks, reducing memory and storage requirements by a factor of 7.

Selected Publications & Patents

Cavia: Camera-controllable Multi-view Video Diffusion with View-Integrated Attention. ICML 2025
Dejia Xu, Yifan Jiang, **Chen Huang**, et al. icml.cc/virtual/2025/poster/44487

3D representation of objects based on a generalized model. US patent application
F. Li, **C. Huang**, H. Tang, S. Pagad, T. Gernoth. US20250316018A1, published October 2025.

Object centric scanning. US patent application
T. Gernoth, **C. Huang**, O. C. Hamsici, S. Feng, H. Tang, T. Rick. US application 17/179,487; publication: September 2021.

Model and ensemble compression for metric learning. US patent
J. M. Susskind, F. Tang, **C. Huang**, S. Sun, W. A. Talbot. US11281993B2, granted March 2022.

Skills

Generative modeling	Diffusion models; image, video, and 3D generation; pretraining, supervised fine-tuning, distillation
3D vision	3D and semantic segmentation, object detection
On-device inference	Post-training quantization, mixed precision, low-bit weights, operation fusion, deployment
Languages & frameworks	Python, C/C++; PyTorch, JAX, TensorFlow, Ray

Education

University of Missouri
Ph.D., 2016
Research: computer vision, object detection, fine-grained recognition

Beihang University
B.S., 2010