# Yicong Hong

Research Scientist at Adobe

**Video Generation & World Models**

www.yiconghong.me mr.yiconghong@gmail.com (+1) (650)-444-5327
345 Park Ave, San Jose, CA 95110, US

## PROJECT HIGHLIGHTS

**RELIC: Interactive Video World Model** – Project Page
- A real-time, generalizable, and interactive video world model with long-horizon spatial memory.

**LRM: Large Reconstruction Model** – Project Page
- A 3D foundation model for creating high-fidelity object mesh from a single image in 5 seconds.

**LaCT: Test-time training done right** – Project Page
- Large-chunk test-time training and scaling memory for NLP, NVS, and AR Video Gen.

**RecBERT: A Recurrent Vision-and-Language BERT** – Project Page
- Build recurrence into multimodal Transformers for sequential decision-making in navigation.

## PROFESSIONAL

**Research Scientist** **Video Generation**
*Adobe Research - Australia → San Jose, United States* *Feb 2024 - Present*
- Interactive video world models, Test-Time Training, linear-complexity video models, and autoregressive long video generation.

**Research Internship** ×2 **3D Reconstruction & Representation Learning**
*Adobe Research - Australia (remote)* *2022 & 2023*
- Large Reconstruction Model (LRM) and egocentric-to-topdown map learning.

**Student Researcher** **Language-Guided Visual Navigation**
*Australian Center for Robotic Vision (ACRV)* *Feb 2019 - Dec 2020*
- Recurrent BERT, graph neural networks, vision-language pre-training.

**Research Internship/Student** **Human Pose and Shape**
*Data61, Commonwealth Scientific and Industrial Research Organisation* *Nov 2017 - Nov 2018*
- 3D human pose and shape visualization from multi-view images.

## EDUCATION

**Ph.D. in Computer Science** **Embodied Vision-and-Language**
*Australian National University* *Feb 2019 - July 2023*
Chair: Prof. Stephen Gould, Supervisors: Prof. Qi Wu, Prof. Lexing Xie

Thesis: Learning Language-Guided Visual Navigation.

**Bachelor of Engineering (First-Class Honours)** **Major in Mechatronic Systems**
*Australian National University* *Feb 2015 - Nov 2018*
GPA: 6.594 / 7

Thesis: Marker-Less Human Pose and Shape Visualisation from Multi-View Images.

## ADDITIONAL PUBLICATIONS

**Pushing the Boundaries of State Space Models for Image and Video Generation** (*preprint 2025*)
- Yicong Hong, Yao Yuan, Long Mai, Feng Liu.
- A State Space Model-Transformer hybrid foundation model for text to image and video generation.

**Progressive autoregressive video diffusion models** (*CVPR 2025*)
- Desai Xie, Zhan Xu, Yicong Hong, Hao Tan, Difan Liu, Feng Liu, Arie Kaufman, Yang Zhou.
- Progressive noise scheduling for long autoregressive video generation.

**Scaling Data Generation in Vision-and-Language Navigation** (*ICCV 2023* ***Oral***)
- Zun Wang, Jialu Li, Yicong Hong*, Yi Wang, Qi Wu, Mohit Bansal, Stephen Gould, Hao Tan, Yu Qiao.
- Generate 4.9 million instruction-trajectory pairs to boost agent's R2R success rate to 80%. *Project Lead

**Bridging the Gap Between Learning in Discrete and Continuous Envs for VLN** (*CVPR 2022*)
- Yicong Hong, Zun Wang, Qi Wu, Stephen Gould.
- A waypoint predictor to enable high-level actions in continuous environments.

**NavGPT: Explicit Reasoning in Vision-and-Language Navigation with LLMs** (*AAAI 2023*)
- Gengze Zhou, Yicong Hong, Qi Wu.
- Use GPT-3.5 and GPT-4 to reason and zero-shot sequential navigation tasks.

**Learning Navigational Visual Representations with Semantic Map Supervision** (*ICCV 2023*)
- Yicong Hong, Yang Zhou, Ruiyi Zhang, Franck Dernoncourt, Trung Bui, Stephen Gould, Hao Tan.
- Encoders trained with contrastive learning between agent's egocentric views and top-down maps.

## OTHERS

### Award

- **1st Place Winner of the (RxR) Habitat Challenge** in Embodied AI workshop, CVPR 2022.
- **Outstanding reviewer** of ICCV2025, WACV 2024, ICCV 2021, CVPR 2021.
- **Outstanding Area Chair** of EMNLP 2024.
- **NVIDIA Academic Hardware Grant 2022**.

### Teaching

**Teaching Assistance** — **Deep Learning (ENGN8536)**
*College of Engineering and Computer Science, ANU* — *2019 & 2020*
- Running labs and tutorials, designing and marking students' assignments, organizing group projects.

**Postgraduate and PhD Student Advisor** — **Vision-and-Language Navigation**
*Australian National University & University of Adelaide* — *2021 - 2024*
- Zun Wang (now PhD at UNC, Chapel Hill) and Gengze Zhou (now PhD at Uni of Adelaide).

### Services

**Reviewer & Program Committee**
- **Reviewer** CVPR'(21-25), ICLR'(22-24), ICCV'(21-25), NeurIPS'(24-25), ICML'25, ICRA'22, AAAI'(21-24), WACV'24, SIGGRAPH'24, EMNLP'(24), ECCV'20, etc.
- **Area Chair** EMNLP'24.

## REFEREES

**Prof. Stephen Gould** - Australian National University - ✉ Email: stephen.gould@anu.edu.au
**Dr. Hao Tan** - Adobe Research - ✉ Email: hatan@adobe.com