Yicong Hong

Ph.D. in Computer Science at CECC, ANU

Vision-and-Language, Text/Image-to-3D Generation, Embodied AI

http://www.yiconghong.me mr.yiconghong@gmail.com (+61) 402 192 382
Brian Anderson Building, 115 North Rd, Acton ACT 2601, Australia

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.

EXPERIENCE

Professional

Research Internship — **Language/Image-to-3D Reconstruction**
Adobe Research - San Jose, United States — *Mar 2023 - Jun 2023*
- Text/Occupancy map guided 3D shape generation via latent diffusion.
- Learning generic representations for single image-to-3D reconstruction.

Research Internship — **Visual Representation Learning**
Adobe Research - San Jose, United States — *Mar 2022 - Nov 2022*
- Representation learning and contrastive learning for visual navigation.

Student Researcher — **Vision-and-Language 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.

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 - 2023*
- Zun Wang: ICCV2023 and CVPR2022 publication in VLN, winner of the RxR Habitat Challenge.
- Gengze Zhou: Applying LLMs in visual navigation.
- Bahram Mohammadi: Applying external knowledge graph for navigation.

Guest Lectures — **Vision-and-Language Research**
College of Engineering and Computer Science, ANU — *2020 - 2021*
- COMP6490 Document Analysis, 2021. & ENGN8536 Deep Learning, 2020.

SELECTED WORKS

LRM: Large Reconstruction Model for Single Image to 3D (*Preprint 2023*)
- Yicong Hong, Kai Zhang, Jiuxiang Gu, Sai Bi, Yang Zhou, Difan Liu, Feng Liu, Kalyan Sunkavalli, Trung Bui, Hao Tan.
- Build transformer-based large vision model to predict 3D shape of an object in an image.

A Recurrent Vision-and-Language BERT for Navigation (*CVPR 2021 Oral*)
- Yicong Hong, Qi Wu, Yuankai Qi, Cristian Rodriguez-Opazo, Stephen Gould.
- Build recurrence into Transformers for sequential decision-making in navigation.

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

Language and Visual Entity Relationship Graph for Agent Navigation (*NeurIPS 2020*)
- Yicong Hong, Cristian Rodriguez-Opazo, Yuankai Qi, Qi Wu, Stephen Gould.
- A graph neural network to model language-object/scene-action correspondence in navigation.

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.

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: Initialized the project and coordinated collaboration among five institutes.

ADDITIONAL

Websites

Open Source Projects https://github.com/YicongHong

Award

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

Services

Reviewer & Program Committee
- CVPR'(23,22,21), ICCV'(23,21), ICLR'(24,23,22), ICRA'22, AAAI'(24,23,22,21), WACV'24, EMNLP'21, SpLU-RoboNLP'21, TNNLS'21, ECCV'20, DICTA'20.

REFEREES

Prof. Stephen Gould - Australian National University - ✉ Email: stephen.gould@anu.edu.au

Prof. Qi Wu - University of Adelaide - ✉ Email: qi.wu01@adelaide.edu.au

Dr. Hao Tan - Adobe Research - ✉ Email: hatan@adobe.com

Prof. Nick Barnes - Australian National University - ✉ Email: nick.barnes@anu.edu.au