

# YUXUAN WANG

Contact: ethan.yuxuan.wang@gmail.com · Homepage: www.yuxuanw.me ·

## EDUCATION

---

**Nanyang Technological University (NTU)**, Computer Science and Engineering,  
Ph.D. Candidate Aug. 2022 - Present

Supervised by Prof. *Hanwang Zhang* in Computer Vision.

**National University of Singapore (NUS)**, Electrical and Computer Engineering,  
Master of Science (*GPA: 4.47/5*) Aug. 2021 - Jun. 2022

Supervised by Prof. *Mike Zheng Shou* in Computer Vision.

**Beihang University**, Electronic Information Engineering,  
Bachelor of Engineering (*GPA: 87.5/100*) Sep. 2016 - Jun. 2020

## PUBLICATIONS

---

### **Nautilus: Locality-aware Autoencoder for Scalable Mesh Generation**

Yuxuan Wang\*, Xuanyu Yi\*, Haohan Weng\*, Xiaokang Wei, Xianghui Yang, Chunchao Guo, Long Chen, Hanwang Zhang

*Arxiv Preprint, Under Review, 2025*

- We propose Nautilus, a locality-aware autoencoder for artist-like mesh generation, which leverages the local properties of manifold meshes to achieve structural fidelity and efficient representation.

### **View-Consistent 3D Editing with Gaussian Splatting**

Yuxuan Wang, Xuanyu Yi, Zike Wu, Na Zhao, Long Chen, Hanwang Zhang

*European Conference on Computer Vision (ECCV), 2024*

- In the diffusion model, we proposed effective multi-view consistency designs that harmonize the inconsistent multi-view image guidance by integrating with 3D Gaussian Splatting (3DGS) characteristics, offering high-quality 3DGS editing.

### **Predicate Debiasing in Vision-Language Models Integration for Scene Graph Generation**

Yuxuan Wang, Xiaoyuan Liu

*Main Conference, Empirical Methods in Natural Language Processing (EMNLP), 2024*

- We introduced a plug-and-play debiasing method for the zero-shot VLMs, dynamically ensembling them to address the underrepresentation issue in Scene Graph Generation (SGG) models.

### **GEB+: A Benchmark for Generic Event Boundary Captioning, Grounding and Retrieval**

Yuxuan Wang, Difei Gao, Licheng Yu, Stan Weixian Lei, Matt Feiszli, Mike Zheng Shou

*European Conference on Computer Vision (ECCV), 2022*

- We introduced three tasks of video boundary understanding on our new dataset called Kinetics-GEB+ (Generic Event Boundary Plus), consisting of over 170k boundaries associated with captions in 12K videos.
- We designed a new Temporal-based Pairwise Difference (TPD) Modeling method for visual difference representation and achieved significant performance improvements.

### **Pushing Rendering Boundaries: Hard Gaussian Splatting**

Qingshan Xu, Jiequan Cui, Xuanyu Yi, Yuxuan Wang, Yuan Zhou, Yew-Soon Ong, Hanwang Zhang

*Arxiv Preprint, Under Review, 2025*

- We propose Hard Gaussian Splatting, dubbed HGS, which considers multi-view significant positional gradients and rendering errors to grow hard Gaussians that fill the gaps of classical Gaussian Splatting on 3D scenes, thus achieving superior NVS results.

## **PBR3DGen: A VLM-guided mesh generation with high-quality PBR texture**

Xiaokang Wei, Bowen Zhang, Xianghui Yang, Yuxuan Wang, Xi Zhao, Yan Luximon

*Arxiv Preprint, Under Review, 2025*

- We present PBR3DGen, a two-stage mesh generation method that produces high-quality PBR materials by integrating a novel multi-view PBR material estimation model and a PBR-based Large Reconstruction Model (PBR-LRM).

## **Symbolic Replay: Scene Graph as Prompt for Continual Learning on VQA Task**

Stan Weixian Lei, Difei Gao, Jay Zhangjie Wu, Yuxuan Wang, Wei Liu, Mengmi Zhang, Mike Zheng Shou

*AAAI Conference on Artificial Intelligence (AAAI), 2023, Oral*

- We introduced Scene Graph as Prompt (SGP) for symbolic replay, a real-data-free replay-based method for Continual Learning VQA, which overcomes the limitations of replay-based methods by leveraging the scene graph as an alternative to images for replay.

## **AssistSR: Task-oriented Video Segment Retrieval for Personal AI Assistant**

Stan Weixian Lei, Difei Gao, Yuxuan Wang, Dongxing Mao, Zihan Liang, Lingmin Ran, Mike Zheng Shou

*Findings, Empirical Methods in Natural Language Processing (EMNLP), 2022*

- We introduce a new dataset and a new task called Affordance-centric Question-driven Video Segment Retrieval (AQVSR), aiming at retrieving affordance-centric instructional video segments given users' questions.
- To address the task, we developed a straightforward model called Dual Multimodal Encoders (DME).

## **INTERNSHIP EXPERIENCE**

---

### **Research Intern | Tencent Hunyuan**

Hunyuan 3D AIGC Center, Tencent TEG

Jun. 2024

### **Software Development Intern | Inspur**

Onsite, Commercial Aircraft Corporation of China

Jan. 2021

AI and Big Data Platform for Health Service

Jun. 2020

## **RESEARCH PROJECTS**

---

### **IMS-MLD Decoding Algorithm for Reed-Muller Code**

Senior Thesis supervised by Prof. *Qin Huang*, Beihang University

Feb. 2020

### **Analog Fountain Code (AFC)**

Supervised by Prof. *Qin Huang*, Beihang University

Jul. 2019

### **Blind Identification and Demodulation of Modulated Signals**

Supervised by Prof. *Qin Huang*, Beihang University

Jul. 2018

## **SKILLS**

---

- **Language:** Mandarin (Native), English (GRE: 331/340, TOEFL: 110/120)
- **Coding:** Python, Java, C#, MATLAB, SQL
- **Music:** Piano, Guitar, Clarinet, Cavalry trumpet, Singing, Chorus Conducting

## **OTHER EXPERIENCE**

---

### **Chorus, Beihang University**

Tenor Singer

Sep. 2016

- Gold medal, National College Student Art Exhibition of China.
- Gold medal, 2018 International Chorus Festival, Calella, Spain.
- Highly acclaimed in concerts and musical theater performances.