

# Quynh Phung

PhD Candidate in Generative AI

+ (240) 643-9418

quynhpt@umd.edu

qqphung.github.io

qqphung

Phung Quynh

HvBv2J3v6Osc

## Education

---

### University of Maryland, College Park

Ph.D. in Computer Science – AI/Computer Vision

GPA: 3.925/4.0, Advisor: Jia-Bin Huang

MD, U.S.A.

2022 – 2026

### Hanoi University of Science and Technology

BEng. in Computer Science – Information Technology

Vietnam

2014 – 2019

## Publications (See Full List)

---

### Trace: Object Motion Editing in Videos with First-Frame Trajectory Guidance

Quynh Phung, Long Mai, Cusuh Ham, Feng Liu, Jia-Bin Huang, Aniruddha Mahapatra

Preprint

### UniVerse: A Unified Modulation Framework for Segmentation-Free Disentangled Multi-Concept Personalization

Quynh Phung, Sandes Ghimire, Minsi Hu, Chung-Chi Tsai, Jia-Bin Huang

CVPR'26

### CineVerse: Consistent Keyframe Synthesis for Cinematic Scene Composition

Quynh Phung, Long Mai, Fabian Caba Heilbron, Feng liu, Jia-Bin Huang, Cusuh Ham

WACV'25

### Grounded Text-to-image Synthesis with Attention Refocusing

Quynh Phung, Songwei Ge, Jia-Bin Huang

CVPR'24

[132 stars & 180 citations](#)

### Zero-Shot Visual Instruction Generation

Quynh Phung, Songwei Ge, Jia-Bin Huang

Preprint

### Explore Image Deblurring via Encoded Blur Kernel Space

Phong Tran, Anh Tran, Quynh Phung, Minh Hoai

CVPR'21

[145 starts & 116 citations](#)

### Toward Realistic Single-View 3D Object Reconstruction with Unsupervised Learning from Multiple Images

Long-Nhat Ho, Anh Tran, Quynh Phung, Minh Hoai

ICCV'21

[48 stars & 12 citations](#)

## Research/ Working Experience

---

### University of Maryland, College Park

Improve number and spacial relationship of text-to-image models: (Advisor: Prof. Jia-Bin Huang) Leveraging the LLMs to for better understanding prompts for text-to-image models. Proposing attention refocusing losses to improve the controllability of text-to-image models, published in CVPR 2024. ([paper](#), [code](#), [project page](#))

Segmentation-Free Text-Guided Multi-Object Personalization: (Advisor: Prof. Jia-Bin Huang) A segmentation-free, text-guided framework for extracting and composing multiple objects from complex scenes for personalized visual generation (CVPR'26).

Generating visual instructions from text instruction: (Advisor: Prof. Jia-Bin Huang) Generating clear, coherent visual instructions from text, ensuring the sequence accurately reflects the instructions while maintaining object consistency. ([paper](#))

### Adobe Inc. – Research Intern

Layout-guidance for video editing: (Advisors: Aniruddha Mahapatra, Dr. Long Mai):

Develop a framework for video editing that modifies an object's trajectory using a user-defined path in the first frame. (*submitted to ECCV'26*)

MD, U.S.A.

02/ 2023 – present

CA, U.S.A

05/2024 – 03/2026

Consistent multi-shot generation: (Advisors: Dr. Cusuh, Dr. Long Mai): A frame-work for movie scene planning by fine-tuning LLMs and using text-to-image models. (1 patent, WACV'25)

## VinAI Research – AI Resident

**Vietnam**

*AI Resident at AI Residency Program* (Supervisor: Dr. Anh Tran)

02/ 2020 – 04/ 2022

3D reconstruction: Unsupervised recovering 3D structure from a single image (ICCV'21).

## Honors & Awards

---

2022-  
present    Reviewer at ICCV, ECCV, WACV, CVPR, SIGGRAPH Asia  
2022        Diversity Fellowship – Ph.D. UMD  
2017        Microsoft YouthSpark Scholarship for Female student in Information Technology

## Skill Sets

---

**Languages:** Python (Proficient), C/C++ (libtorch), TensorRT.

**Frameworks:** PyTorch, libtorch, Tensorflow (1.0, 2.0), TensorRT, Scikit-Learn, Numpy, OpenCV, huggingface