Jing Tan

CONTACT · sparkstj.github.io · jingtan@ie.cuhk.edu.hk · +86 15050863237

RESEARCH INTERESTS
3D Scene Creation, 3D-aware Image and Video Synthesis, World Model
Video Understanding

EDUCATION

Department of Information Engineering, The Chinese University of Hong Kong Hong Kong SAR
Ph.D. Student in MMLab Aug. 2023 – Present
Supervisor: Prof. Dahua Lin

Department of Computer Science and Technology, Nanjing University Nanjing, China
Master in MCG Lab Sept. 2020 – June 2023
Supervisor: Prof. Limin Wang

Kuang Yaming Honors School, Nanjing University Nanjing, China
B.Sc., Major in Computer Science and Technology (GPA 4.44/5, Rank 3/29). Sept. 2016 – June 2020

PUBLICATION

- **Imagine360: Immersive 360 Video Generation from Perspective Anchor**,
  Jing Tan*, Shuai Yang*, Tong Wu, Jingwen He, Yuwei Guo, Ziwei Liu, Dahua Lin, (*: equal contribution)
  NeurIPS, 2025.
  Imagine360 is the *first* framework that lifts standard perspective video into 360-degree video with rich and structured motion, unlocking dynamic scene experience from full 360 degrees. [Project Page]
- **LayerPano3D: Layered 3D Panorama for Hyper-Immersive Scene Generation**, [Github ★ 292]
  Shuai Yang*, **Jing Tan***, Mengchen Zhang, Tong Wu, Yixuan Li, Gordon Wetzstein, Ziwei Liu, Dahua Lin, (*: equal contribution)
  SIGGRAPH (conference track), 2025.
  LayerPano3D generates full-view, explorable panoramic 3D scene from a single text prompt.[Project Page]
- **Temporal Perceiver: A General Architecture for Generic Boundary Detection**,
  Jing Tan, Yuhong Wang, Gangshan Wu, Limin Wang,
  T-PAMI, 2023.
  Temporal Perceiver is a general architecture based on Transformer decoders, that offers a unified solution to detect arbitrary generic boundaries, including shot-level, event-level and scene-level temporal boundaries. [Project Page]
- **PointTAD: Multi-label Temporal Action Detection with Learnable Query Points**,
  Jing Tan, Xiaotong Zhao, Xintian Shi, Bin Kang, Limin Wang,
  NeurIPS, 2022.
  PointTAD effectively tackles multi-label TAD by introducing a set of learnable query points to represent the action keyframes. [Project Page]
- **Relaxed Transformer Decoders for Direct Action Proposal Generation**, [Cites: 255]
  Jing Tan*, Jiaqi Tang*, Limin Wang, Gangshan Wu,
  ICCV, 2021. (*: equal contribution)
  The first transformer-based framework for temporal action proposal generation. [Project Page]

- **SS4D: Native 4D Generative Model via Structured Spacetime Latents**,
  Zhibing Li, Mengchen Zhang, Tong Wu, **Jing Tan**, Jiaqi Wang, Dahua Lin,
  SIGGRAPH Asia (TOG), 2025.
  A native 4D generative model that synthesizes dynamic 3D objects directly from monocular video. [Project Page]
- **GenDoP: Auto-regressive Camera Trajectory Generation as a Director of Photography** ,
  Mengchen Zhang, Tong Wu, **Jing Tan**, Ziwei Liu, Gordon Wetzstein, Dahua Lin,
  ICCV, 2025.
  An auto-regressive model that generate artistic and expressive camera trajectories from text prompts and geometric cues. [Project Page]
- **IDArb: Intrinsic Decomposition for Arbitrary Number of Input Views and Illuminations**,
  Zhibing Li, Tong Wu, **Jing Tan**, Mengchen Zhang, Jiaqi Wang, Dahua Lin,
  ICLR, 2025.
  IDArb is a Diffusion-based intrinsic decomposition framework for an arbitrary number of image inputs under varying illuminations. [Project Page]
- **HumanVid: Demystifying Training Data for Camera-controllable Human Image Animation**,
  Zhenzhi Wang, Yixuan Li, Yanhong Zeng, Youqing Fang, Yuwei Guo, Wenran Liu, **Jing Tan**, Kai Chen,

Tianfan Xue, Bo Dai, Dahua Lin,
NeurIPS D&B Track, 2024.
Camera-controllable human image animation dataset and framework. [Project Page]

- **Dual DETRs for Multi-Label Temporal Action Detection**,
Yuhan Zhu, Guozhen Zhang, **Jing Tan**, Gangshan Wu, Limin Wang,
CVPR, 2024.
A new Dual-level query-based TAD framework to precisely detect actions from both instance-level and boundary-level.

Experience

AWS AI Lab June 2025 - Nov. 2025
Applied Scientist Intern, Manager: Yantao Shen Bellevue, WA, US

• I'm very fortunate to collaborate with scholars: Prof. Zhuowen Tu and Prof. Jiajun Wu on my intern project, focusing on RL-guided Spatial-aware Image Editing.
• With the help of RLVR, we aim to achieve precise geometric object transformation in images following text instructions.

Tencent, PCG Dec. 2021 - March 2023
Research Intern with Beijing, China
• Worked on Multi-label Video Action Detection via Learnable Query Points. Build a more general video action detector that locates a specific action within a number of simultaneous actions.

Honors and Awards (selected)

• CSIG Master's Thesis Incentive Program Awardee	China Society of Image and Graphics	2025
• Outstanding Master Thesis Award (3/226)	Nanjing University	2024
• Outstanding Graduate of Nanjing University (20%)	Nanjing University	2023
• National Scholarship (6/226)	Ministry of Education	2022
• 3rd Place, GEBD-LOVEU Challenge @ CVPR'22	Workshop Committee	2022
• Outstanding Postgraduate Student Award	Nanjing University	2021

Services

Regular Reviewer for CVPR, ICCV, ECCV, NeurIPS, ICLR and AAAI.

Teaching Assistant
IERG 4998 & 4999 Final Year Project (2023-2024) CUHK
Digital Logic Design and Computer Organization (2021 Spring) Nanjing University