Ziyang Song

+86 177 4855 6311 | songzy305@gmail.com | https://szy-young.github.io

EDUCATION

The Hong Kong Polytechnic University — Hong Kong SAR, China

PhD in Computing; Supervisor: Bo Yang — Sep. 2021 – Dec. 2025

- **Thesis title:** Segmentation and Reconstruction of 3D Objects from Dynamic Scenes

Xi'an Jiaotong University — Xi'an, China

MEng in Control Science and Engineering — Sep. 2018 – Jun. 2021

Xi'an Jiaotong University — Xi'an, China

BEng in Automation (Honors Youth Program) — Sep. 2014 – Jun. 2018

WORK EXPERIENCE

TikTok — San Jose, USA

Research Intern; Mentor: Xinyu Gong — Jul. 2025 – Nov. 2025

- Research in reference-to-video generation
  - Development of a new video generation algorithm with **3D consistency to multi-view reference** images of objects
  - Curation of a large-scale multi-view reference-to-video dataset

SenseTime — Shanghai, China

Research Intern; Mentor: Dongliang Wang — Feb. 2021 – Jul. 2021

- Research in 3D human motion synthesis (*ActFormer*, ICCV 2023)
- Development of a sparse-view 3D human motion capture system

Tencent Robotics X — Shenzhen, China

Research Intern; Mentor: Wanchao Chi — Jun. 2019 – Aug. 2019

- Development of a real-time human action recognition system on NVIDIA Xavier mobile platform

SELECTED PUBLICATIONS

- **Ziyang Song**, Xinyu Gong, Bangya Liu, et al. *MV-S2V: Multi-View Subject-Consistent Video Generation*. Under Review.
- **Ziyang Song**, Jinxi Li, Zihui Zhang, et al. *Mono360: Free View Synthesis of Deformable Objects from Monocular Videos via Dynamic Fusion*. Under Review.
- **Ziyang Song**, Jinxi Li, Bo Yang. *OSN: Infinite Representations of Dynamic 3D Scenes from Monocular Videos*. ICML, 2024.

- **Ziyang Song**, Bo Yang. *Unsupervised 3D Object Segmentation of Point Clouds by Geometry Consistency*. TPAMI, 2024.
- Liang Xu*, **Ziyang Song***, Dongliang Wang, et al. *ActFormer: A GAN-based Transformer towards General Action-Conditioned 3D Human Motion Generation*. ICCV, 2023.
- **Ziyang Song**, Bo Yang. *OGC: Unsupervised 3D Object Segmentation from Rigid Dynamics of Point Clouds*. NeurIPS, 2022.
- Zelin Zhao, Xinyu Gong, **Ziyang Song**, et al. *CETCam: Camera-Controllable Video Generation via Consistent and Extensible Tokenization*. Under Review.
- Peng Huang, Jinxi Li, **Ziyang Song**, et al. *CoVIP: Visual-Physical 3D Scene Reconstruction from a Single Image*. Under Review.
- Jinxi Li, **Ziyang Song**, Bo Yang. *TRACE: Learning 3D Gaussian Physical Dynamics from Multi-view Videos*. ICCV, 2025.
- Jinxi Li, **Ziyang Song**, Siyuan Zhou, et al. *FreeGave: 3D Physics Learning from Dynamic Videos by Gaussian Velocity*. CVPR, 2025.
- Jinxi Li, **Ziyang Song**, Bo Yang. *NVFi: Neural Velocity Fields for 3D Physics Learning from Dynamic Videos*. NeurIPS, 2023.

(* denotes equal contribution)

TECHNICAL SKILLS

- **Frameworks:** PyTorch (deep learning), Taichi (differentiable high-performance geometry processing), Nvdiffrast (differentiable surface rendering)
- **Large-scale GPU usage:** Multi-node distributed training experience of video generation models