

摄影测量大模型

Matrix3D: Large Photogrammetry Model All-in-One

姚遥

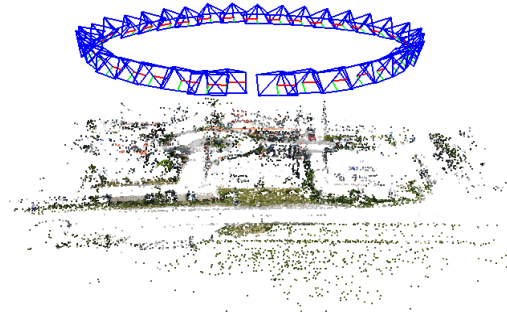
南京大学 NJU-3DV

Introduction | Photogrammetry Pipeline



(a) Multi-view images

SfM



(b) Camera parameters

MVS



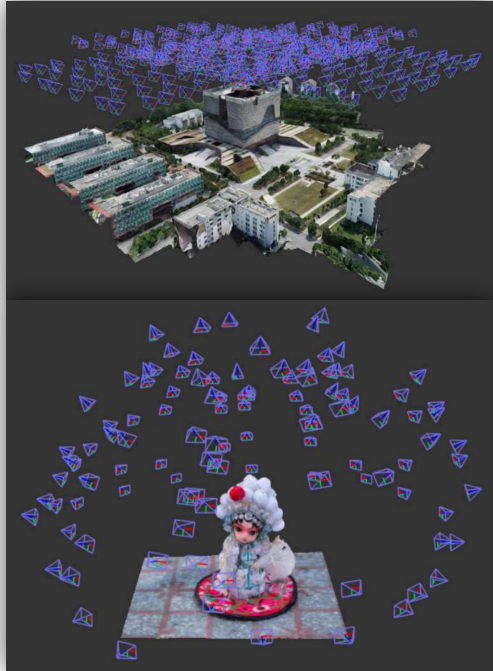
(c) Dense point cloud



NVS
Meshing
...

Designing a Unified Model for 3D Reconstruction

- Applicable to multiple sub-tasks (e.g., SfM, MVS, NVS)



Reconstruction

- Colmap
- MVSNet
- NeRF
- 3DGS
- Dust3r
- ...

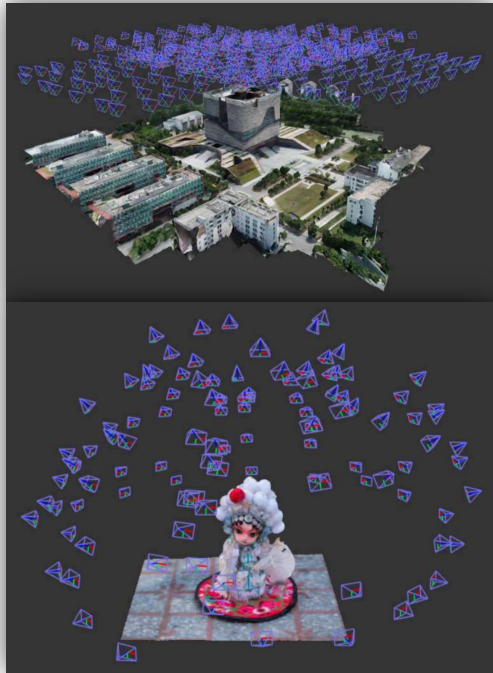
Multi-view ←

Observations

Designing a Unified Model for 3D Content Creation

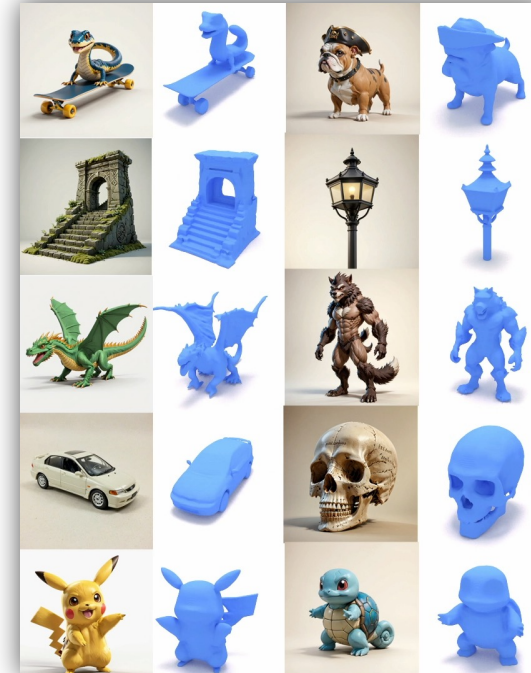
- Applicable to multiple sub-tasks (e.g., SfM, MVS, NVS)

Introduction | 3D Content Creation



Reconstruction

- Colmap
- MVSNet
- NeRF
- 3DGS
- Dust3r
- ...



Generation

- Clay
- Direct3D
- CraftsMan
- TripoSG
- Trellis
- ...

Multi-view



Observations

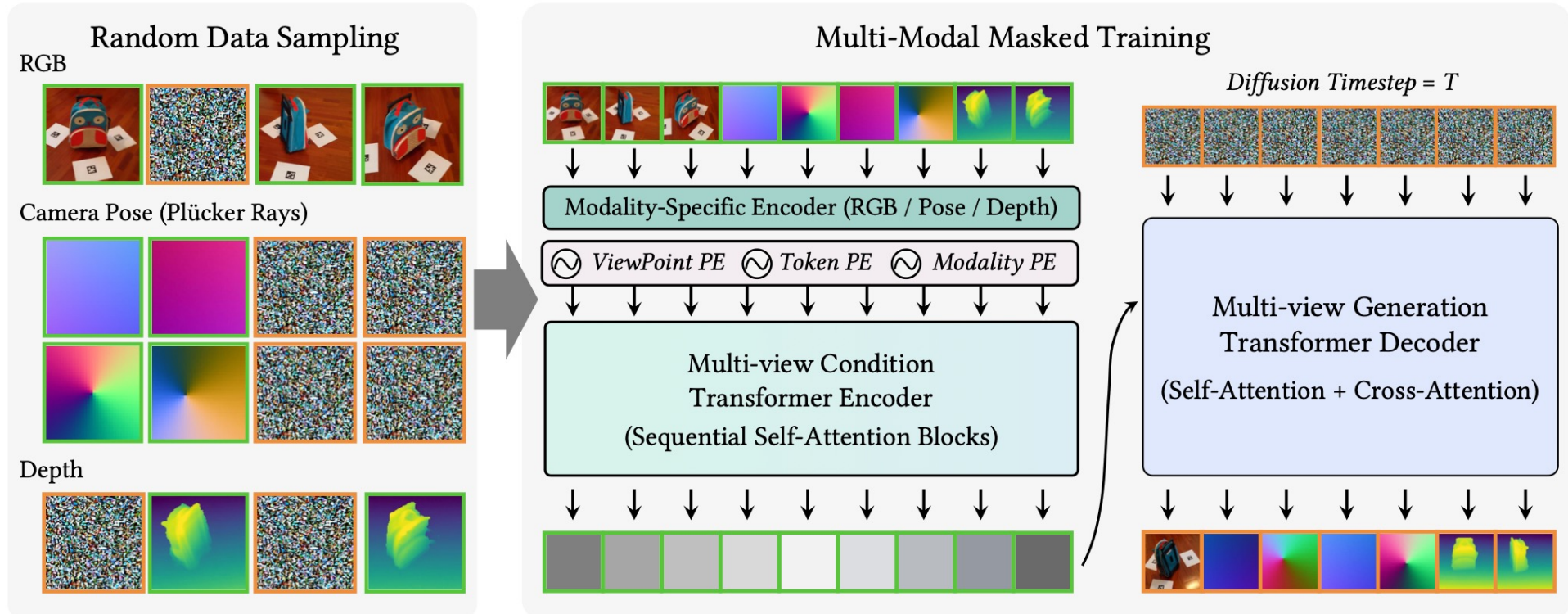


Single-view

Designing a Unified Model for 3D Content Creation

- Applicable to multiple sub-tasks (e.g., SfM, MVS, NVS)
- For both reconstruction and generation

Matrix3D | Large Photogrammetry Model All-in-One



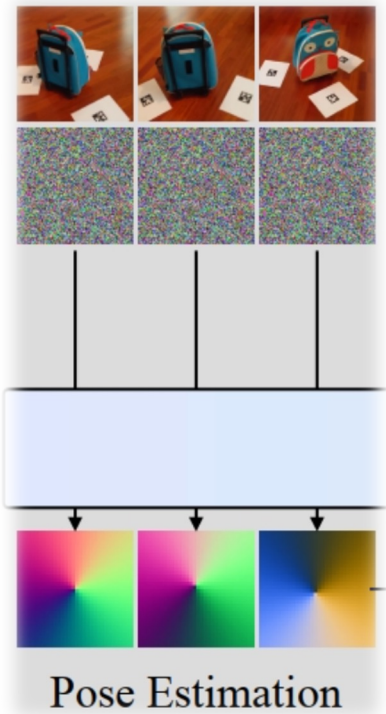
Generative Architectures:

- Multi-modal multi-view Diffusion Transformer
- Photogrammetry tasks as modality conversion
- Mask learning for multi-task training

Unified Representations

- Images: RGB images
- Cameras: Ray maps
- Geometry: depth maps

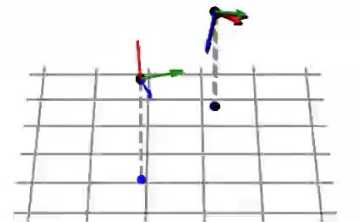
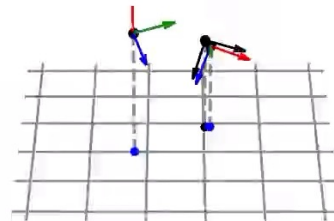
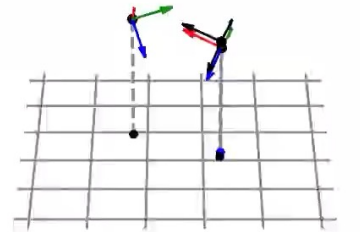
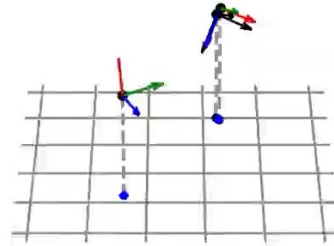
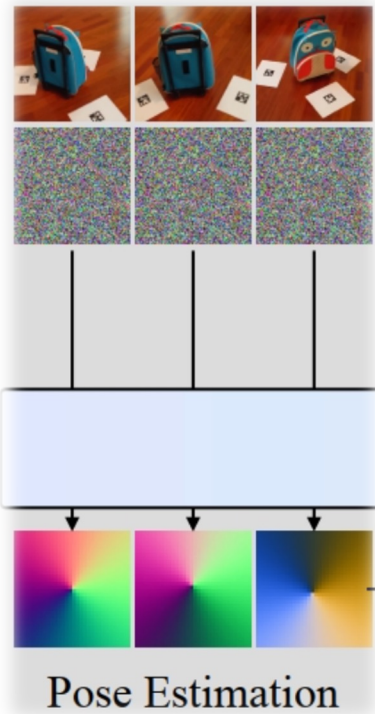
Matrix3D | Pose Estimation



Input Setting

- Images
- Noisy cameras

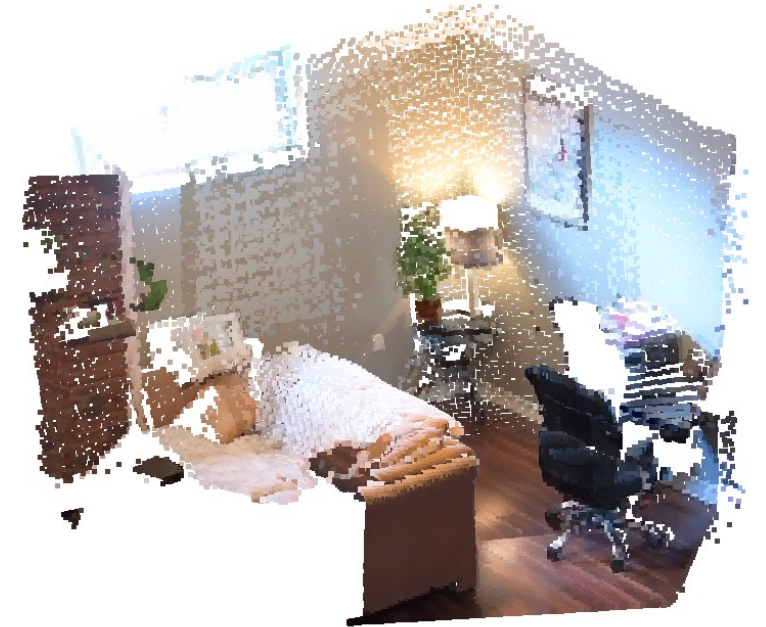
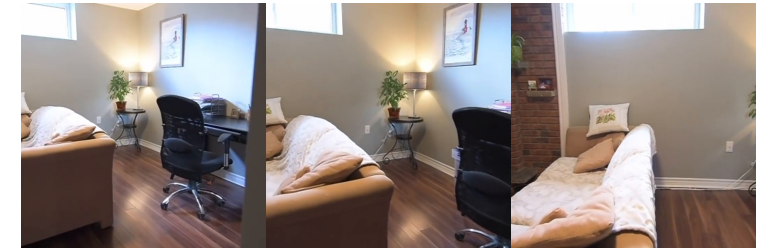
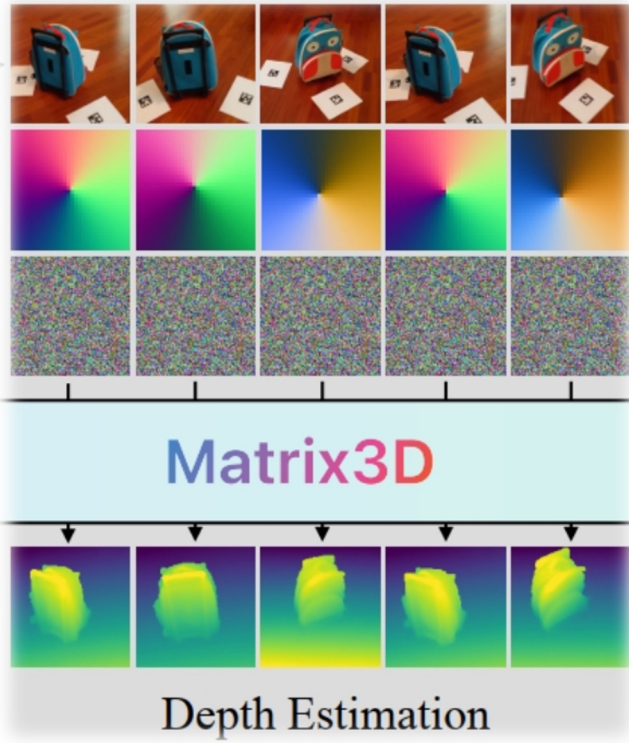
Matrix3D | Pose Estimation



Input Setting

- Images
- Noisy cameras

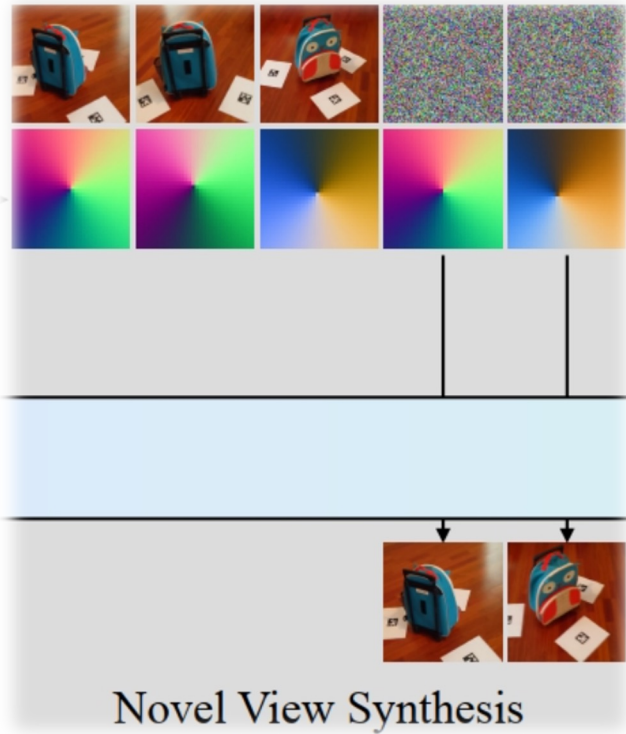
Matrix3D | Multi-view Depth Estimation



Input Setting

- Images
- Cameras
- Noisy depth maps

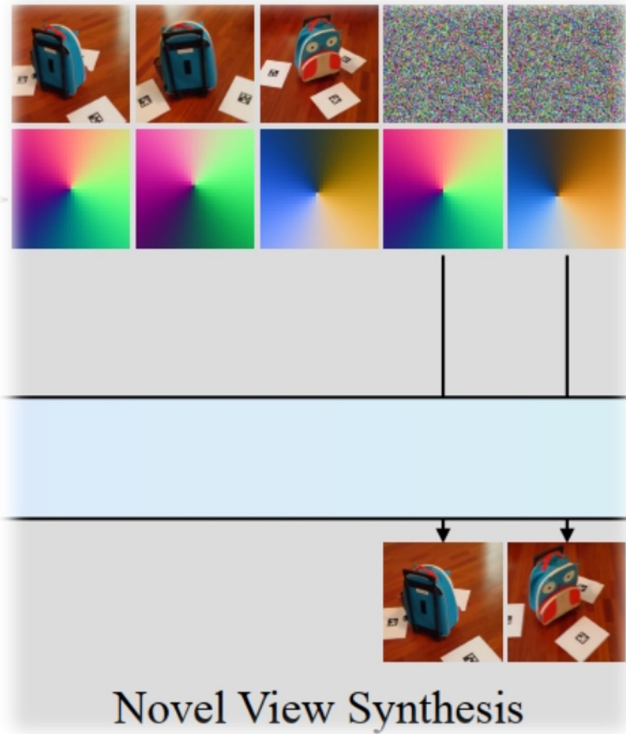
Matrix3D | Novel View Synthesis



Input Setting

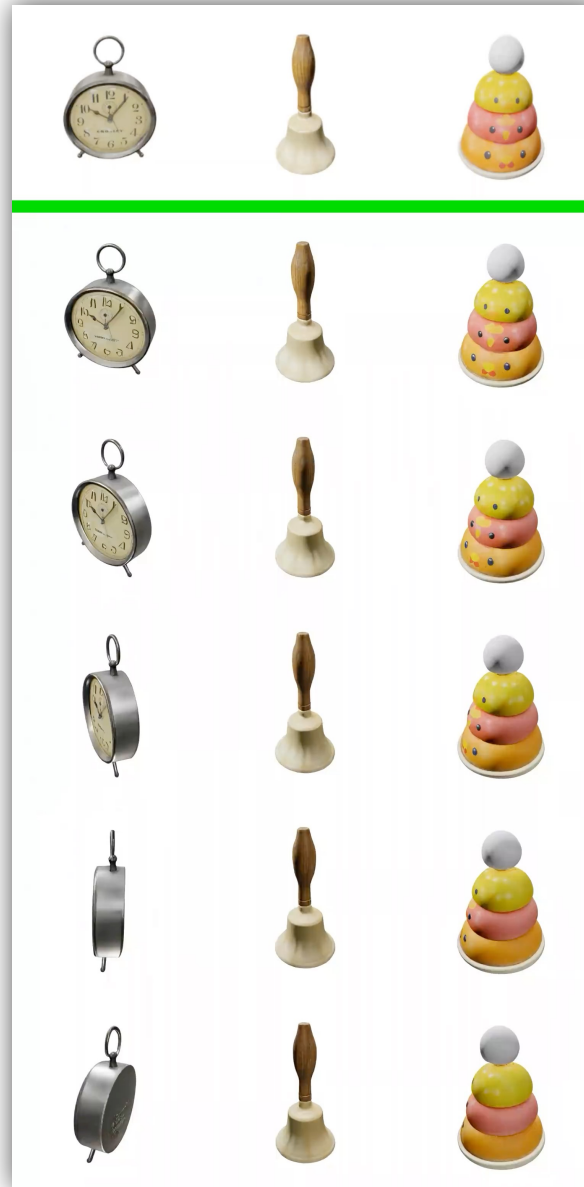
- Images
- Cameras
- Novel cams and noisy images

Matrix3D | Novel View Synthesis

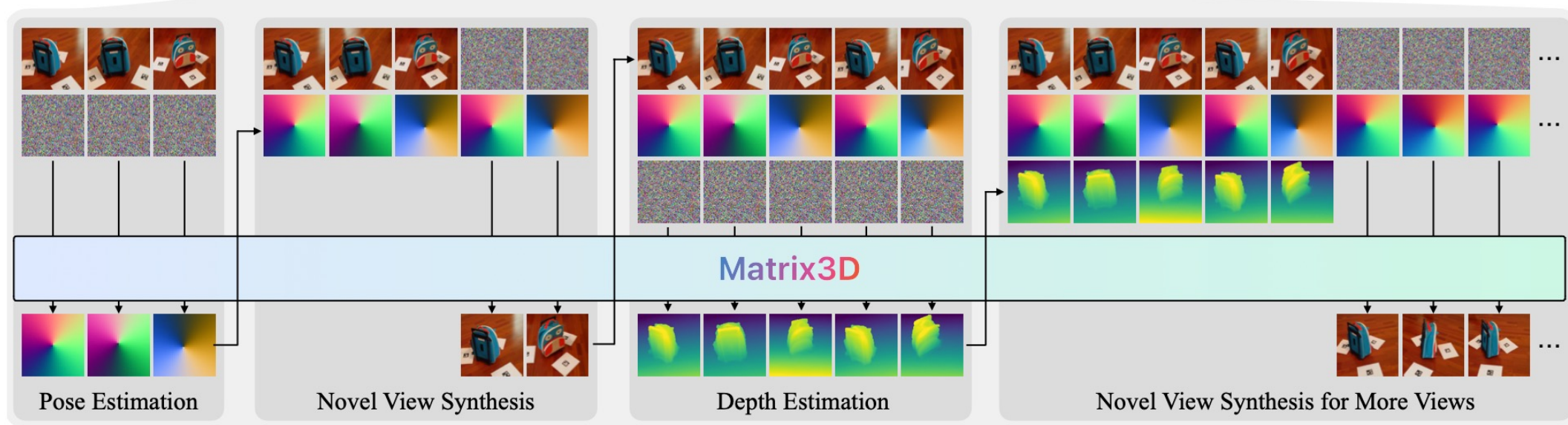
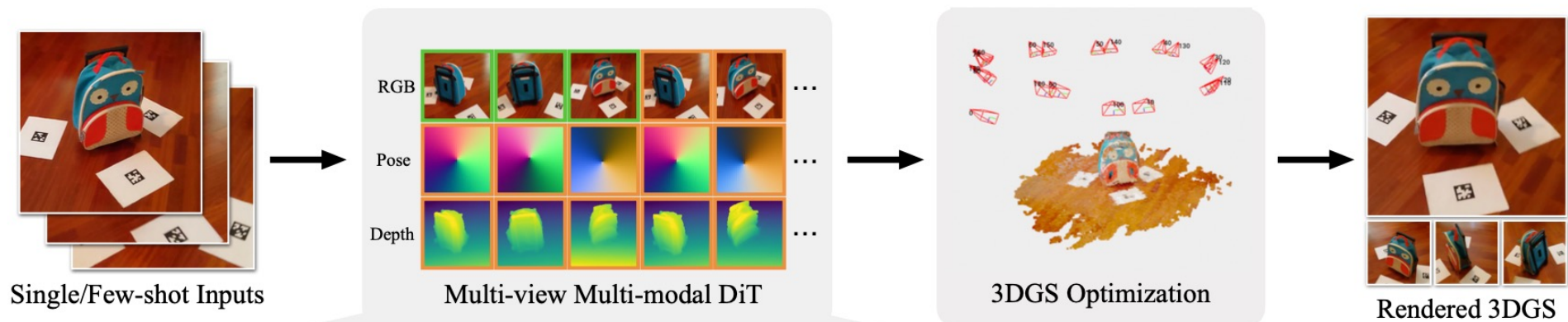


Input Setting

- Images
- Cameras
- Novel cams and noisy images



Matrix3D | Unposed Few-shot Reconstruction



One model, the whole pipeline

Matrix3D | Unposed Few-shot Reconstruction



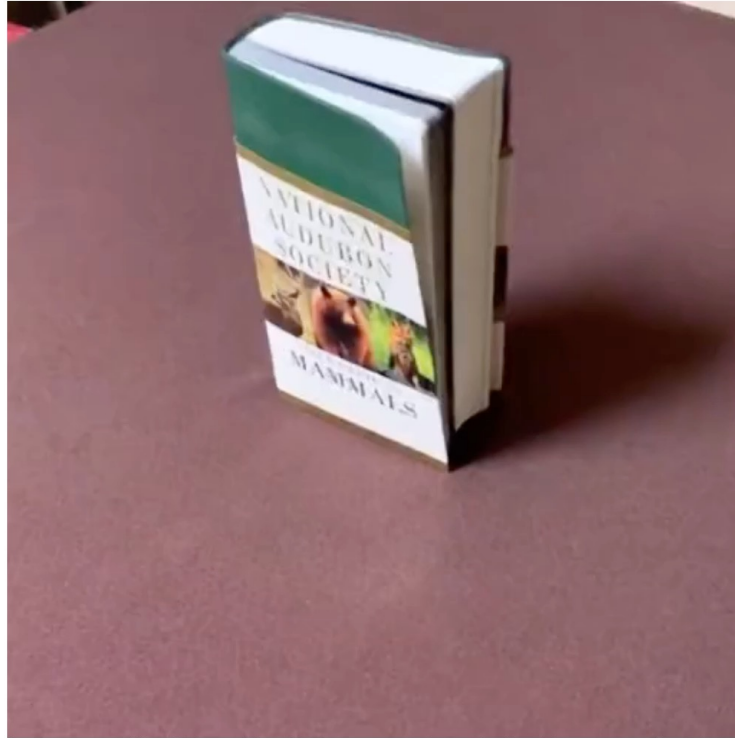
Unposed Input



Back-Projected Point Cloud



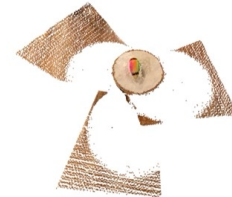
3DGS Optimization



3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings

Matrix3D | Unposed Few-shot Reconstruction



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



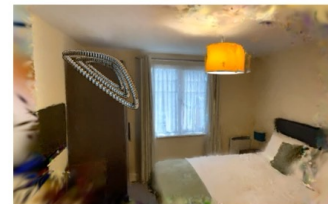
3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings

Matrix3D | Unposed Few-shot Reconstruction



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings



Unposed Input



Back-Projected Point Cloud



3DGS Optimization



3DGS Renderings

Matrix3D: Large Photogrammetry Model All-in-One

CVPR 2025 **(Highlight)**

Yuanxun Lu^{1*}, Jingyang Zhang^{2*}, Tian Fang², Jean-Daniel Nahmias², Yanghai Tsin², Long Quan³,
Xun Cao¹, Yao Yao¹✉, Shiwei Li²

¹Nanjing University, ²Apple, ³The Hong Kong University of Science and Technology

*Equal Contribution ✉Corresponding Author

[Paper](#)

[arXiv](#)

[Code \(coming soon\)](#)

TL;DR: We present **Matrix3D**, a **unified model** that performs several photogrammetry subtasks, including pose estimation, depth prediction, and novel view synthesis using **the same model**.

