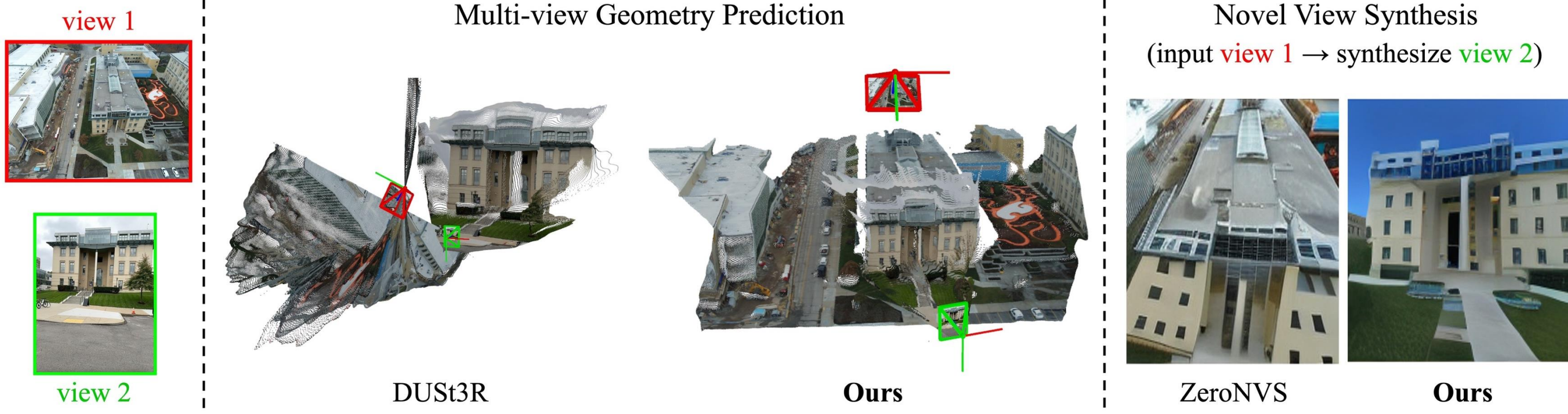




AerialMegaDepth: Learning Aerial-Ground Reconstruction and View Synthesis

Problem setup + Overview

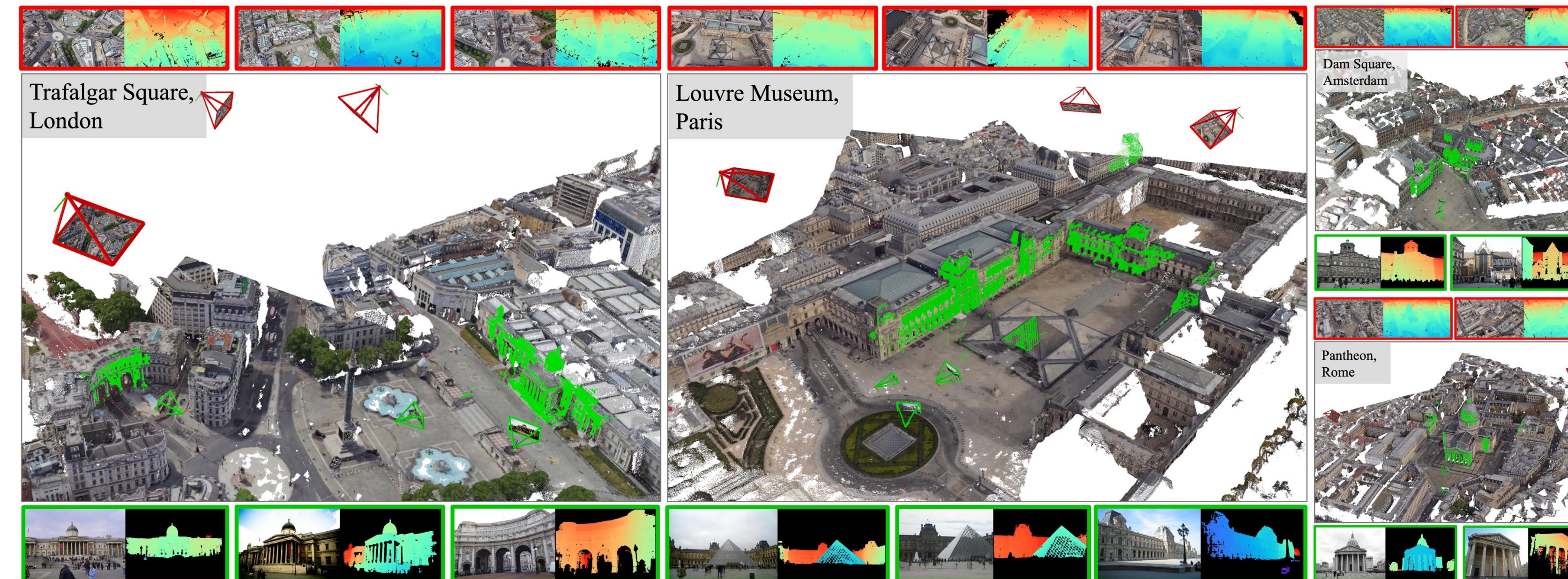
Goal: aerial-ground 3D reconstruction + view synthesis



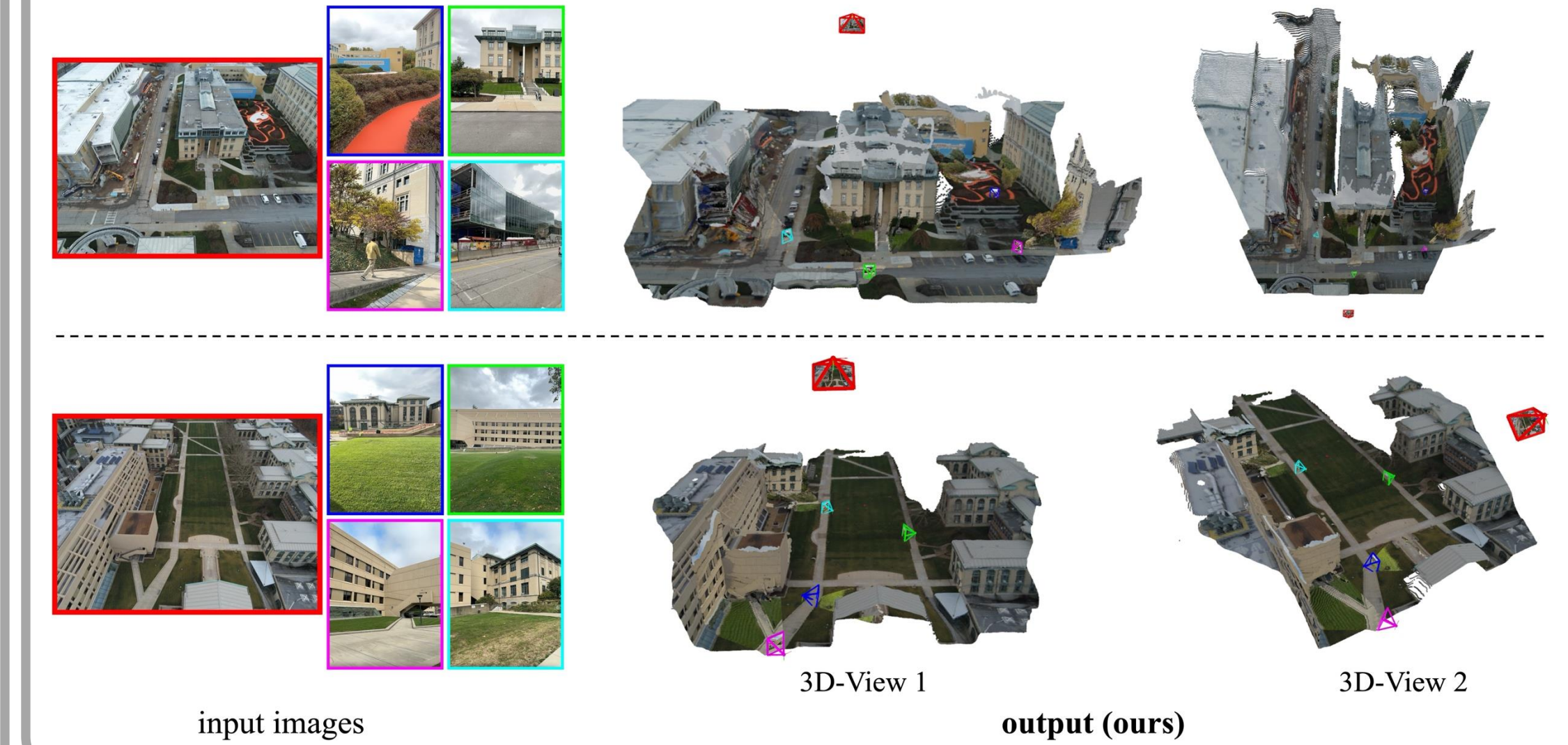
Issue: the lack of real aerial-ground 3D supervision data

AerialMegaDepth: cross-view geometry data

A hybrid varying-altitude 3D dataset combining MegaDepth images with geospatial mesh renderings, featuring 132K images across 137 scenes with camera intrinsics, poses, and depths in a unified coordinate frame



Aerial image helps register ground images

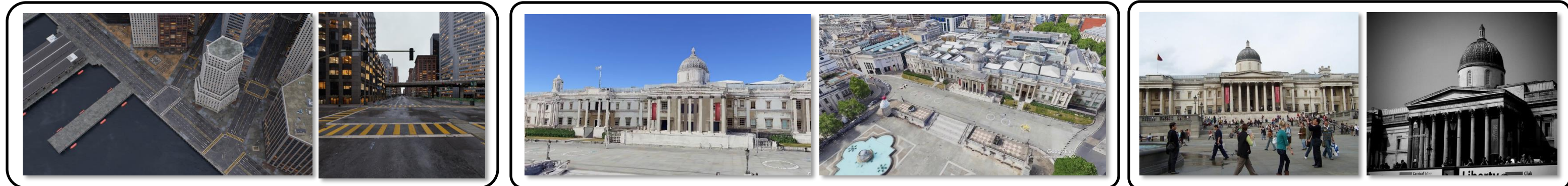


Automatically generating hybrid aerial-ground 3D data

Synthetic data

Pseudo-synthetic data

Real data



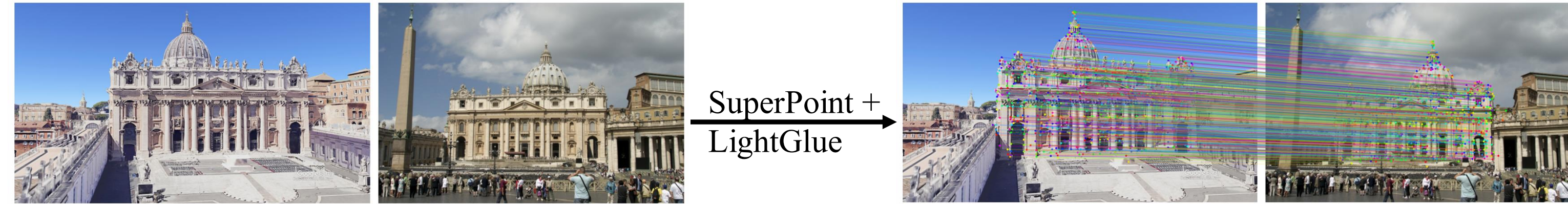
- ✓ extreme viewpoint
- ✗ large sim2real gap
- ✗ not easy to scale

- ✓ extreme viewpoint
- ✗ ~ some domain gap

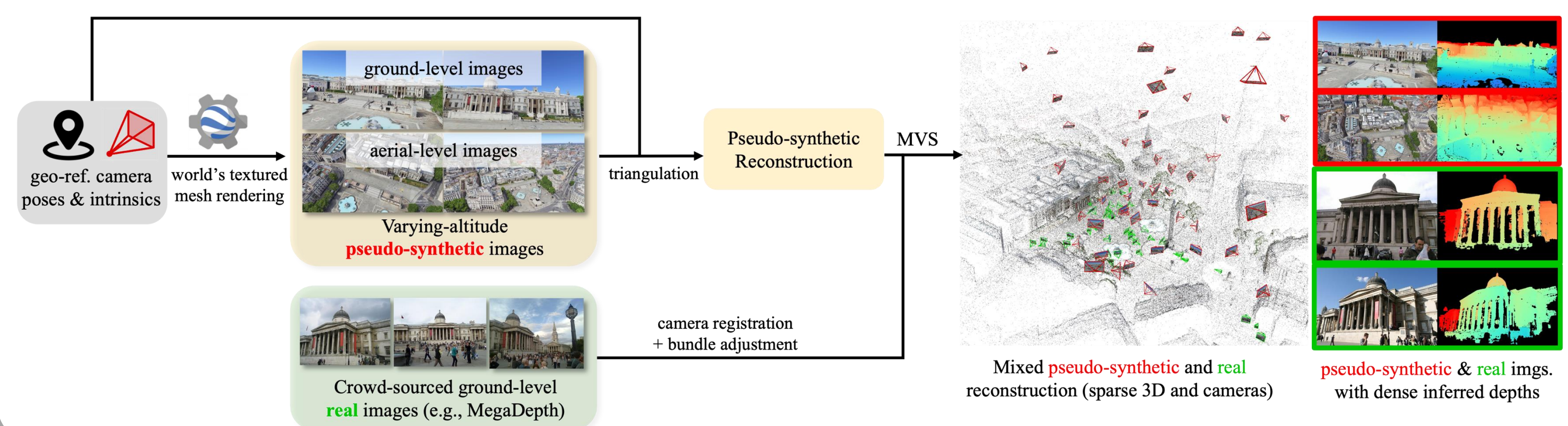
- ✗ no extreme viewpoint
- ✓ diverse & scalable

hybrid data (ours): a combination!

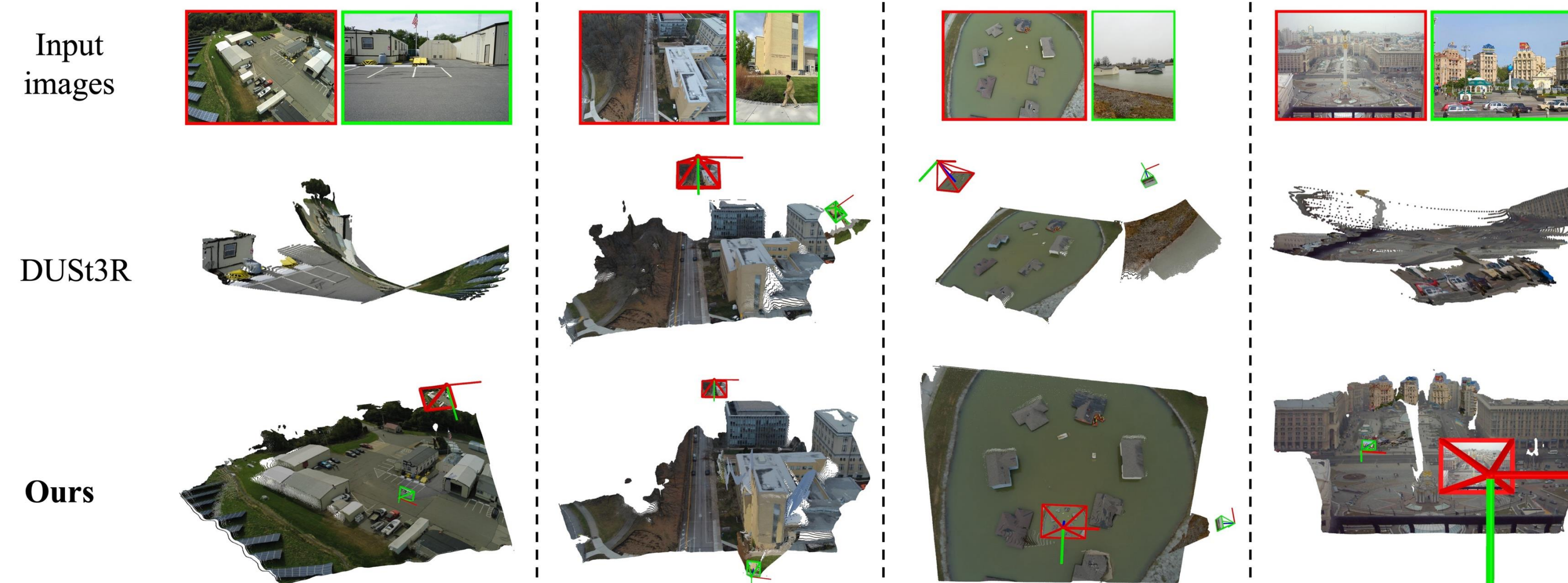
Pseudo-synthetic image Real image



Our data generation framework: pseudo-synthetic renderings + real ground images

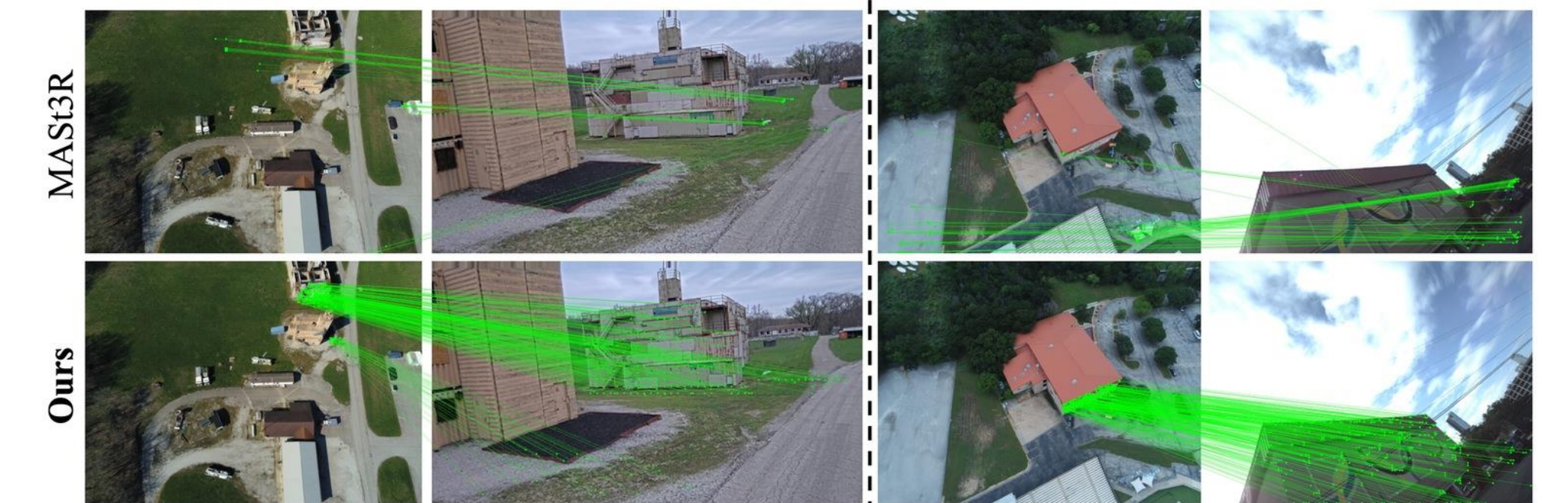


Improving aerial-ground pose + geometry estimation



Additional results

Robust correspondences across ground-aerial pairs



Method	Cam. Rotation Acc.			Cam. Translation Acc.		
	@5°	@10°	@15°	@5°	@10°	@15°
DUST3R (released)	5.20	7.95	9.48	2.75	5.81	9.17
DUST3R + MatrixCity	17.85	37.28	42.80	11.33	25.24	33.24
DUST3R + PSynth (Ours)	31.28	47.63	51.61	28.78	45.66	51.47
DUST3R + Hybrid (Ours)	55.96	71.25	76.15	46.48	68.20	72.78

Improving aerial-to-ground novel view synthesis

