Leveraging Vision Reconstruction Pipelines for Satellite Imagery

Kai Zhang, Jin Sun, Noah Snavely (Cornell University / Cornell Tech)
https://kai-46.github.io/VisSat/

Motivation

Remote Sensing Community

- Expertise in satellite images
- Two-view-based methods

Computer Vision Community

- Expertise in ground-level perspective images
- Multi-view-based methods

Our contribution: adapt vision pipelines to satellite imagery

Typical Vision Reconstruction Pipeline

A set of images → SfM → MVS → Dense point cloud

SfM outputs:
- Camera poses
- Sparse point cloud

Challenges

- Rational polynomial coefficients (RPC) cameras
- Average scene depth is far bigger than scene depth variation

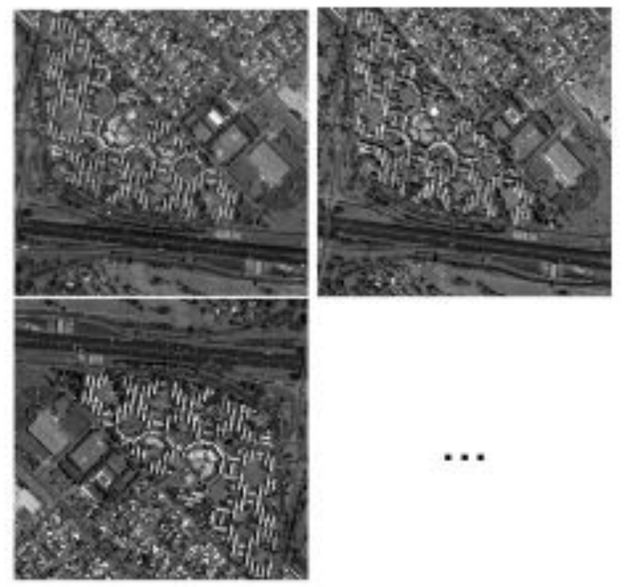
$$row = \frac{P_1(lat, lon, alt)}{P_2(lat, lon, alt)}$$

$$col = \frac{P_3(lat, lon, alt)}{P_4(lat, lon, alt)}$$

- P_1, P_2, P_3, P_4 are four cubic polynomials
- 78 coefficients
- Difficult to vectorize inverse projection & triangulation
- Challenging to harness GPUs

Our Approach

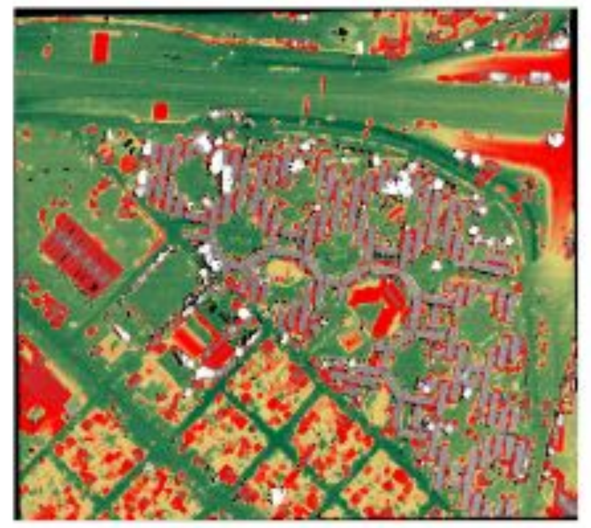
Input: satellite imagery, initial RPC cameras, AOI specification



SfM:
- Image cropping & Tone mapping & RPC correction
- 3D bounding cube generation
- Perspective approximation of RPC cameras
- Bundle adjustment
- Skew correction

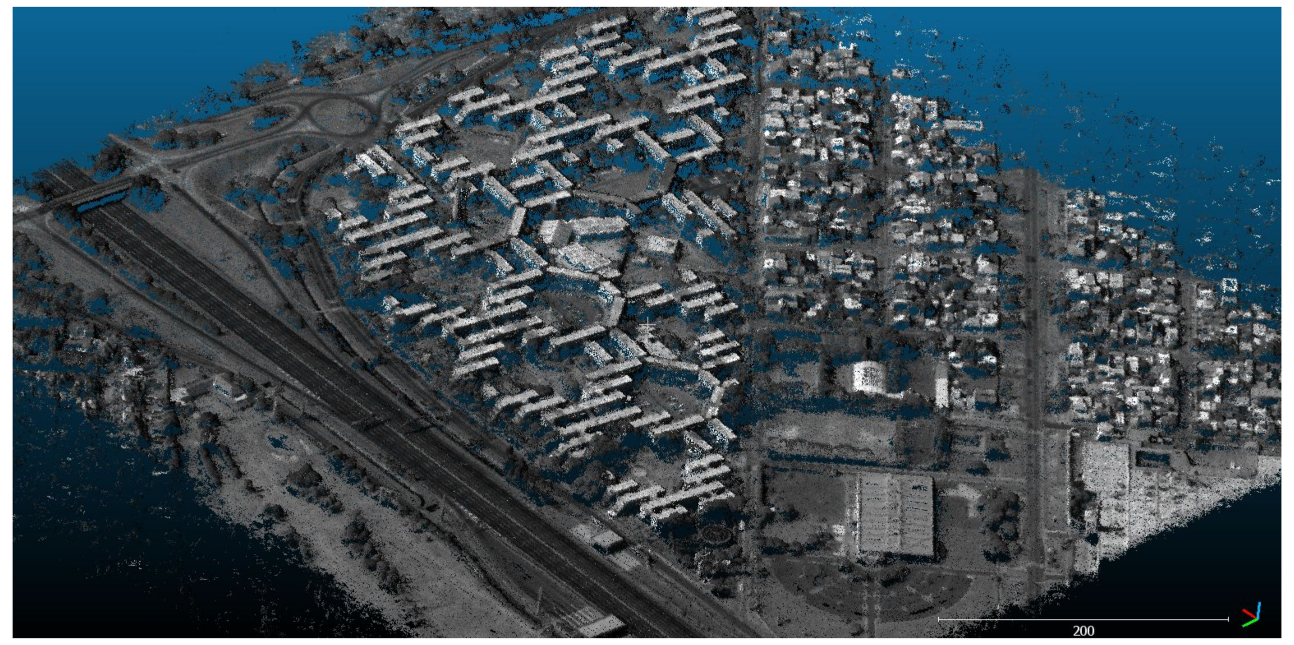
MVS with depth reparametrization

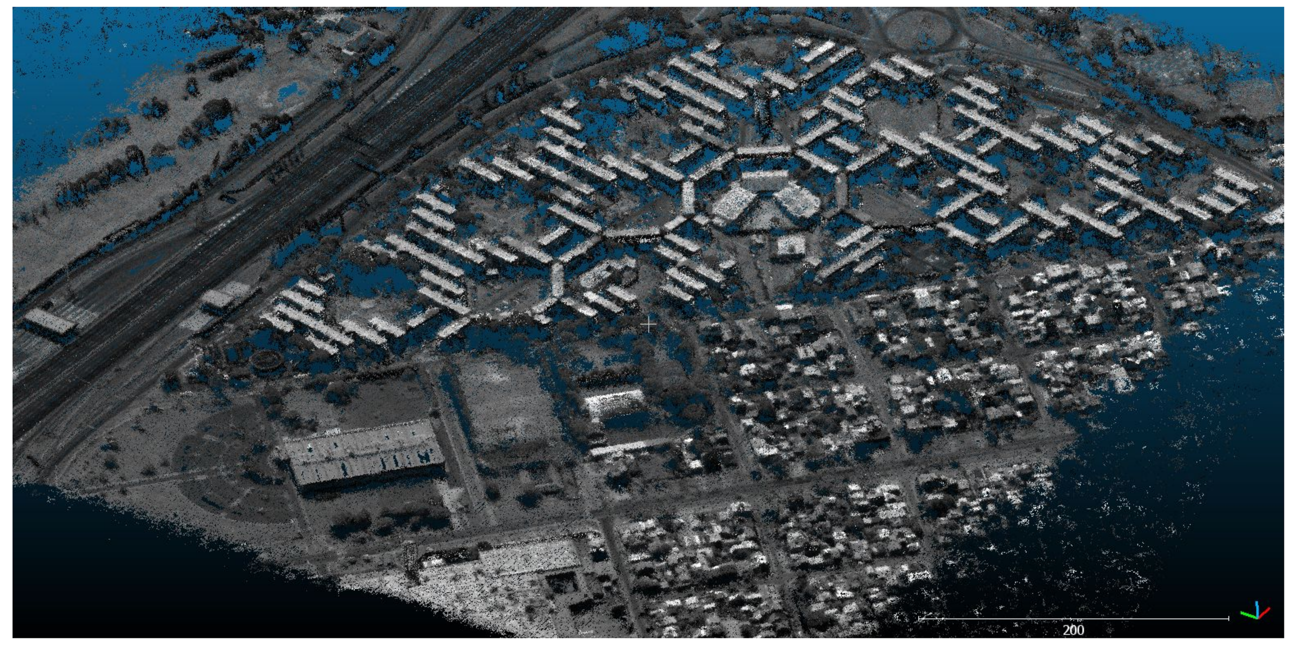
Output: dense point cloud reconstruction



Example Result by Adapted COLMAP

Argentina, 47 images





Data & Evaluation

- MVS3DM benchmark: public satellite MVS dataset
- Flatten the reconstructed point cloud into a height map, then compare with the ground-truth height map

Table 2. Metric result of COLMAP MVS; CP stands for completeness, ME for median height error.

	CP (%)	ME (m)	Time (mins)
site 1	72.5	0.315	18.7
site 2	66.8	0.450	7.37
site 3	63.4	0.393	6.94

Table 3. Comparison between different algorithms on stereo pairs; for each site, both metrics and time are averaged over 10 selected pairs on which S2P performs very well. PSS: plane sweep stereo. CVF: cost volume filtering. MRF: Markov random field.

		CP (%)	ME (m)	Time (mins)
Site 1	S2P	65.6	0.432	16.8
	COLMAP	58.0	0.648	0.374
	PSS+CVF	68.1	0.442	4.40
	PSS+CVF+MRF	**69.1**	**0.395**	29.7
Site 2	S2P	62.8	**0.435**	2.63
	COLMAP	59.3	0.689	0.196
	PSS+CVF	63.8	0.502	1.19
	PSS+CVF+MRF	**64.0**	0.478	6.75
Site 3	S2P	53.9	**0.421**	2.78
	COLMAP	42.1	0.949	0.208
	PSS+CVF	57.1	0.591	2.39
	PSS+CVF+MRF	**58.8**	0.521	14.8

Advantages

- State-of-the-art multi-view stereo solver
- Scalable to GPU