SUPPLEMENT: CUMULATIVE ASSESSMENT FOR URBAN 3D MODELING

Shea Hagstrom, Hee Won Pak, Stephanie Ku, Sean Wang, and Myron Brown

JHU Applied Physics Laboratory

ABSTRACT

In this supplement, we provide metric results for industry capabilities that are more mature than our current open source baseline. Deficiencies shown in these results better reflect unsolved challenges in the current state of the art. With modest baseline improvements, particularly in semantic segmentation, we hope public leaderboard performance will soon be competitive with these results to encourage more attention to improvements in model fitting.

1. INDUSTRY RESULTS

In Tab. 1, we present metric results for the open source baseline both with and without model fitting and compare to three industry capabilities - one with model fitting and two without. Industry capabilities are de-identified and labeled A, B, and C.

Semantic segmentation: Industry solutions considered here have good semantic segmentation performance for these three U.S. cities. Similarly, [1] reports IOU_c of 0.86 for our Jacksonville test site and 0.82 for San Diego. We expect that replacing the baseline semantic segmentation solution with a high-performing public SpaceNet solution [2] will achieve comparable performance for these cities which are reasonably similar to those in the SpaceNet corpus.

MVS: Industry MVS solutions perform well for these test sites for which dozens of cloudless, geometrically diverse satellite images are available. In comparison, the open source baseline MVS point clouds are noisier as reflected in the RMS_z metrics. The bilateral filter [3] we applied to VisSat [4] point clouds in post-processing did improve RMS_z, but we expect more significant improvements to require revisions in the COLMAP pipeline [5]. An alternate MVS pipeline designed specifically for satellite images is S2P [6] which is also open source and was the winning solution from the 2016 Multi-View Stereo 3D Mapping challenge [7].

Model Fitting: Model fitting results for both the open source baseline and industry solution demonstrate improvement in roof slope accuracy for the large flat roofs observed in downtown Jacksonville; however, roof slopes from MVS alone are more accurate for the San Diego test site where roofs are more complex and roof pitches are often higher. We hope that the cumulative metrics we have presented will inspire research to more robustly address these challenges.

	IOU_c	IOU_z	IOU_m	RMS_z	RMS_θ
JACKSONVILLE					
Baseline MVS	0.67	0.56	0.41	0.50	8.80
Baseline Model	0.68	0.56	0.50	0.71	3.53
Industry A MVS	0.83	0.73	0.62	**0.37**	4.58
Industry B MVS	**0.84**	**0.74**	0.63	0.39	4.36
Industry C Model	**0.84**	0.71	**0.67**	0.59	**2.18**
SAN DIEGO					
Baseline MVS	0.58	0.47	0.32	0.70	15.89
Baseline Model	0.59	0.47	0.38	0.99	10.32
Industry A MVS	0.76	**0.67**	**0.51**	**0.52**	**7.88**
Industry B MVS	**0.77**	0.66	0.49	0.61	9.29
Industry C Model	**0.77**	0.62	**0.51**	1.00	8.66
OMAHA					
Baseline MVS	0.58	0.52	0.34	0.37	9.32
Baseline Model	0.58	0.50	0.40	0.57	8.35
Industry A MVS	**0.92**	**0.87**	**0.72**	**0.25**	4.53
Industry B MVS	0.87	0.81	0.64	0.28	5.54
Industry C Model	0.86	0.77	0.66	0.50	**3.62**

Table 1. IOU_m and RMS_θ metrics show the value of roof modeling for improving 3D reconstruction.

2. ACKNOWLEDGMENTS

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA) contract no. 2017-17032700004. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA or the U.S. Government.

3. REFERENCES

[1] X. Li, Y. Lin, J. Miller, A. Cheon, and W. Dixon, “Primitive-based 3d building modeling, sensor simulation, and estimation,” in *IGARSS 2019*, 2019, pp. 5148–5151.

[2] A. V. Etten, Dave Lindenbaum, and Todd M. Bacastow,

"Spacenet: A remote sensing dataset and challenge series," *ArXiv*, vol. abs/1807.01232, 2018.

[3] Julie Digne and Carlo de Franchis, "The Bilateral Filter for Point Clouds," *Image Processing On Line*, vol. 7, pp. 278–287, 2017.

[4] Kai Zhang, Jin Sun, and Noah Snavely, "Leveraging Vision Reconstruction Pipelines for Satellite Imagery," in *ICCV Workshop on 3D Reconstruction in the Wild (3DRW)*, 2019.

[5] Johannes Lutz Schönberger and Jan-Michael Frahm, "Structure-from-Motion Revisited," in *CVPR*, 2016.

[6] G. Facciolo, C. De Franchis, and E. Meinhardt-Llopis, "Automatic 3d reconstruction from multi-date satellite images," in *CVPRW*, 2017, pp. 1542–1551.

[7] M. Bosch, Z. Kurtz, S. Hagstrom, and M. Brown, "A multiple view stereo benchmark for satellite imagery," in *IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2016.