STPLS3D: A Large-Scale Synthetic and Real Aerial Photogrammetry 3D Point Cloud Dataset

Meida Chen[1], Qingyong Hu[2], Zifan Yu[3], Hugues Thomas[4], Andrew Feng[1], Yu Hou[5], Kyle McCullough[1], Fengbo Ren[3], Lucio Soibelman[1]

[1]University of Southern California, [2]University of Oxford, [3]Arizona State University, [4]University of Toronto, [5]Carnegie Mellon University

Motivation and Objective

❖ Deep learning algorithms are data-hungry, especially in the 3D domain.

❖ This research aims to investigate the possibility of using synthetic photogrammetric data to augment/substitute real-world data for training 3D point cloud segmentation algorithms.

Synthetic data generation pipeline

Input: GIS data (i.e., DSM, building footprints, road vectors.)
Output: synthetic photogrammetric 3D point cloud with annotations.

Synthetic 3D point cloud with annotations
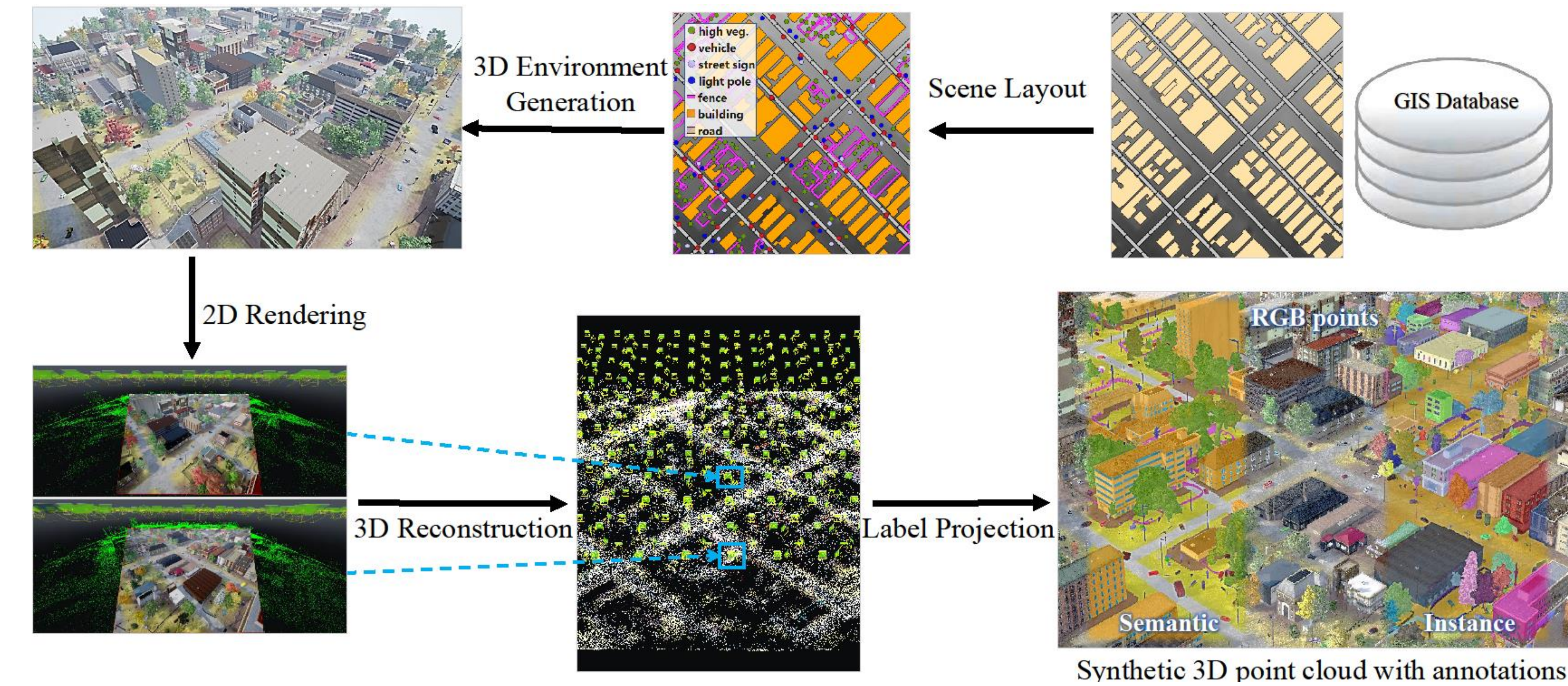
✓ Procedurally generating scene layouts.
✓ Procedurally generating highly detailed 3D building models.
✓ Using a large game object and material database.
✓ Simulating aerial image collections following real-world setups.
✓ Photogrammetric 3D reconstruction using rendered images.

Experiments and results

Semantic segmentation (real vs. synthetic vs. synthetic+real)

Training sets	Methods	mIoU (%)	oAcc (%)	Per Class IoU (%)					
				Ground	Building	Tree	Car	Light pole	Fence
Real subsets	PointTransformer [84]	49.40	85.85	85.23	47.77	**76.72**	39.51	28.61	18.56
	RandLA-Net [35]	51.84	84.79	88.14	46.88	61.40	48.72	46.04	19.83
	SCF-Net [22]	53.79	86.66	**89.19**	53.12	65.28	**48.91**	46.59	19.63
	MinkowskiNet [17]	52.85	83.28	82.76	40.30	71.68	47.00	49.33	26.04
	KPConv [72]	**57.80**	**87.20**	86.69	**63.41**	66.32	46.36	**56.08**	**27.95**
Synthetic subsets	PointTransformer [84]	58.65	92.01	90.42	**74.54**	85.18	31.76	42.36	27.67
	RandLA-Net [35]	59.38	91.33	90.15	69.20	82.21	50.13	40.36	24.20
	SCF-Net [22]	58.82	90.49	89.53	62.39	81.55	**52.99**	44.10	22.36
	MinkowskiNet [17]	56.17	90.55	90.74	66.11	78.63	36.86	36.41	**28.26**
	KPConv [72]	**61.92**	**92.35**	**91.41**	68.31	**86.00**	48.97	51.99	24.82
Real+Synthetic	PointTransformer [84]	62.14	91.96	89.74	**74.79**	84.73	45.10	46.75	**31.72**
	RandLA-Net [35]	61.38	92.31	91.25	68.71	84.35	55.04	43.30	23.83
	SCF-Net [22]	61.89	92.10	90.99	68.69	84.99	55.58	45.36	25.71
	MinkowskiNet [17]	62.59	**93.16**	91.66	74.70	**87.97**	48.80	43.95	28.49
	KPConv [72]	**65.01**	93.03	**91.86**	71.44	87.12	54.77	**55.39**	29.48

Instance segmentation baselines

	Metric	mean (%)	Build.	Low Veg.	Medium Veg.	High Veg.	Vehicle	Truck	Aircraft	Military Veh.	Bike	Motorcycle	Light Pole	Street Sign	Clutter	Fence
HAIS [16]	AP	**35.1**	66.8	20.9	17.6	23.2	75.7	51.9	42.6	31.1	7.4	50.8	47.0	8.3	22.6	25.7
	AP50	**46.7**	73.9	35.7	25.0	29.2	86.9	61.3	65.2	39.2	17.0	69.0	62.9	13.7	27.9	46.5
	AP25	**52.8**	75.9	46.8	31.9	32.1	89.0	66.0	72.0	44.5	22.1	75.4	68.1	15.0	31.7	68.4
PointGroup [41]	AP	23.3	60.0	11.6	10.7	19.2	58.7	39.8	27.6	21.2	2.2	12.0	23.7	8.1	13.9	18.1
	AP50	38.5	70.4	28.3	19.0	25.4	83.9	57.9	47.9	35.3	7.9	44.0	46.8	14.7	19.6	38.4
	AP25	48.6	73.7	43.8	23.7	29.5	87.9	61.4	59.8	42.3	19.4	68.1	66.8	16.6	22.6	64.9

Released datasets

62 Synthetic datasets:
- 46,281 rendered images
- 16 km² coverage
- Up to 18 semantic labels with instance annotations

4 real-world datasets:
- 16,376 aerial images
- 1.27 km² coverage
- 6 semantic labels
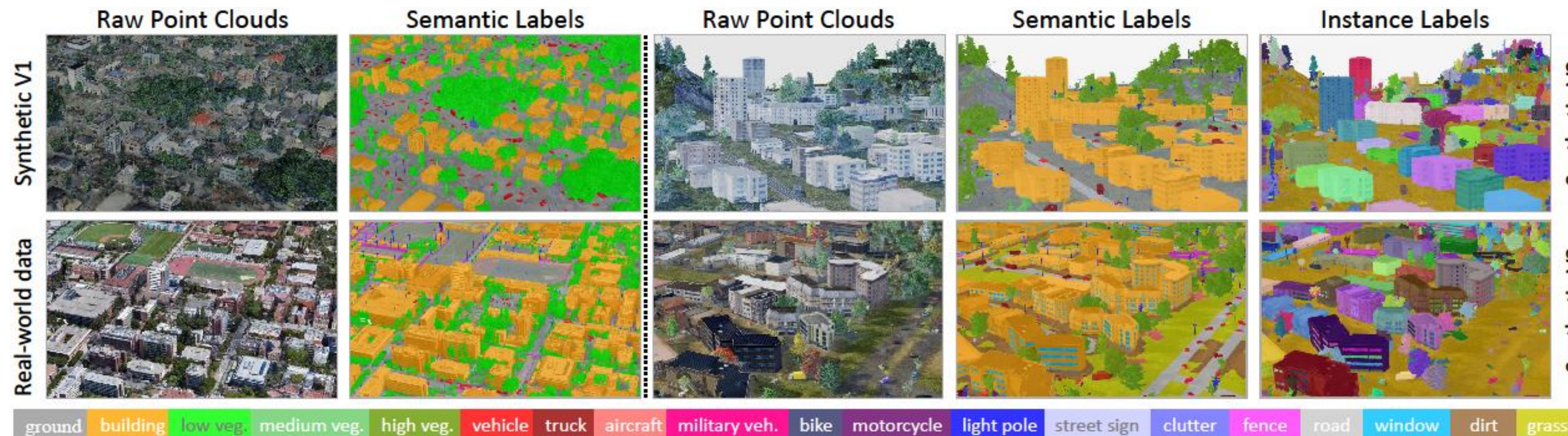
Links

Project page: www.stpls3d.com

Code

CodaLab Evaluation server

Urban3D 2022 Workshop

Demo