

Supplementary Material: Efficient 2D and 3D Facade Segmentation using Auto-Context

Raghudeep Gadde*, Varun Jampani*, Renaud Marlet, and Peter Gehler



In this supplementary, we present more qualitative and class-wise quantitative results on the facade segmentation datasets described in the main paper. Figure 1 shows visual results from different stages of auto-context classifier and also the result obtained by applying a pairwise Potts model and grammar based prior. Tables 1, 2, 3, 4 and 5 show a leader-board containing class-wise performance of relevant state-of-the-art methods and three stages of the proposed auto-context based technique on the eTRIMS [1], CMP [2], Graz [3], LabelMeFacades [4] and ENPC Art-deco [5] datasets respectively. Similarly, class-wise results on the RueMonge2014 [6] dataset for the tasks of image segmentation, point cloud segmentation and mesh segmentation are provided in Tables 6, 7 and 8 respectively. Finally, we provide more visual results which have been selected based on the absolute overall pixel-accuracy of ST3 and include images with the (i) highest, (ii) average, and (iii) lowest performance in Figures 2, 3, 4, 5, 6 and 7.

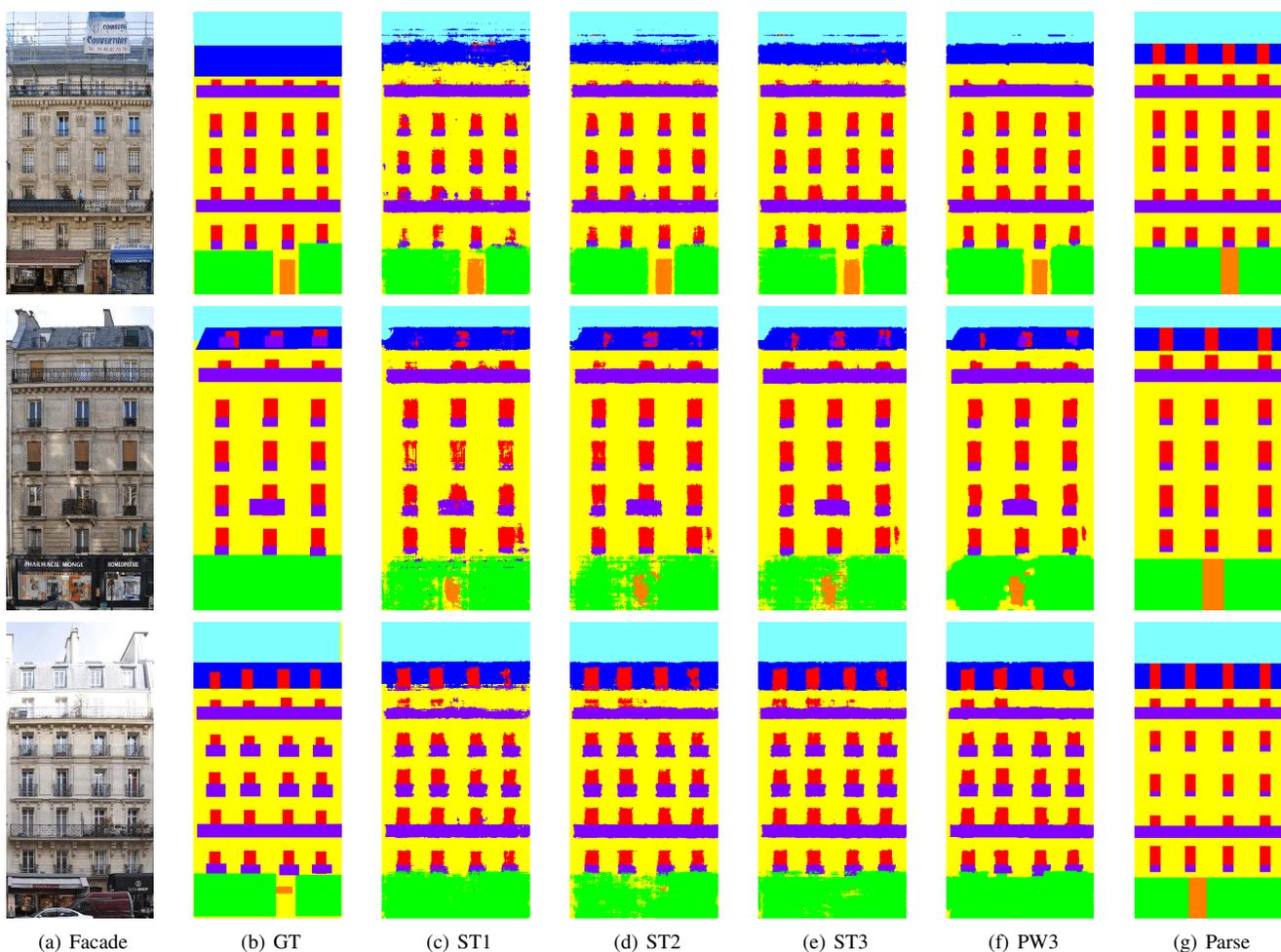


Fig. 1. (a) Sample facade images from ECP dataset; (b) Ground truth segmentation; and (c,d,e) Result of various classification stages of our auto-context method. Observe that the method removes isolated predictions and recovers the second lowest line of windows. (f) Potts model on top of ST3 result, and (g) parsed result obtained by applying reinforcement learning [7] using ST3 result.

- Raghudeep Gadde and Renaud Marlet are with LIGM (UMR 8049), Ecole des Ponts, UPE, Champs-sur-Marne, France.
- Varun Jampani is with MPI-IS, Germany.
- Peter Gehler is with BCCN, University of Tübingen and MPI-IS, Germany.
- The first two authors have contributed equally to this work.

Class	[8]	[9]	[10]	[11]	Auto Context (AC)			AC + Potts Model		
					ST1	ST2	ST3	PW1	PW2	PW3
Building	91	91	84	92	90.3	90.5	90.9	92.7	92.5	92.5
Car	69	70	51	70	63.3	74.8	72.4	69.4	79.1	76.6
Door	18	18	73	20	62.7	62.3	63.6	66.0	63.6	65.3
Pavement	33	33	55	33	43.0	46.5	47.1	43.1	48.6	48.8
Road	55	57	81	56	78.2	82.3	80.3	80.9	84.7	82.1
Sky	93	97	99	96	97.6	98.5	98.6	98.2	98.8	98.9
Vegetation	89	90	92	91	91.1	92.1	92.3	92.4	92.8	92.9
Window	74	71	78	70	65.9	67.1	68.4	65.6	66.5	68.2
Average	65.3	65.9	66.4	66	74.01	76.78	76.7	76.04	78.32	78.14
Overall	83.16	83.84	83.40	83.5	84.68	85.95	86.12	86.39	87.29	87.29
IoU	-	-	-	-	58.7	61.26	61.48	61.49	63.39	63.54

Table 1. Segmentation results of various methods on eTRIMS dataset.

Class	[2]	Auto Context			AC + Potts Model		
		ST1	ST2	ST3	PW1	PW2	PW3
Background	58	67.1	71.8	72.6	68.0	72.6	73.1
Facade	73	74.6	75.3	75.2	80.5	79.9	79.3
Window	61	71.6	76.1	77.0	74.1	77.4	78.1
Door	54	37.9	45.5	47.0	39.6	46.4	48.7
Cornice	41	39.1	47.5	49.6	40.0	48.3	50.1
Sill	27	21.1	32.8	36.2	16.9	30.3	34.6
Balcony	46	31.6	44.1	46.7	31.6	45.2	48.1
Blind	48	22.7	35.8	40.1	19.5	34.7	39.9
Deco	24	10.4	13	13.8	6.1	10.0	11.4
Molding	54	63.2	65.4	66.5	64.2	66.0	67.2
Pillar	25	5.71	11.2	13.6	1.33	7.72	9.78
Shop	59	40.9	45.6	45.6	42.8	46.7	46.8
Average	47.5	40.50	47.00	48.65	40.38	47.1	48.92
Overall	60.3	61.83	65.47	66.24	64.46	67.48	68.08
IoU	-	29.26	34.46	35.86	30.67	36.02	37.47

Table 2. Segmentation results of various methods on CMP dataset.

Class	Auto Context (AC)			AC + Potts Model			[3]	[11]
	ST1	ST2	ST3	PW1	PW2	PW3		
Door	57.3	62.4	62.7	57.3	62.8	63	41	60
Window	78.2	81.2	81.5	77.8	80.6	80.9	60	84
Wall	94.9	94.7	94.9	95.8	95.6	95.8	84	96
Sky	87.4	91.2	90.5	87.7	91.4	90.6	91	93
Average	79.47	82.40	82.42	79.65	82.61	82.56	69	83.5
Overall	90.18	91.02	91.16	90.78	91.53	91.68	78	92.5
IoU	71.25	73.31	73.25	72.49	74.45	74.39	58	-

Table 3. Segmentation results of various methods on Graz dataset.

Class	[12]	[13]	Auto Context(AC)			AC + Potts Model		
			ST1	ST2	ST3	PW1	PW2	PW3
Building	-	-	87.7	88.1	88.2	92.7	91.8	92.1
Car	-	-	47.1	53.6	54.8	51.1	57.0	58.2
Door	-	-	6.52	6.03	5.12	2.61	3.22	1.71
Pavement	-	-	24	25.3	24.6	22.0	24.2	23.3
Road	-	-	80.3	82.1	84.5	85.3	85.1	87.6
Sky	-	-	86.2	87.2	87.4	88.3	88.6	88.9
Vegetation	-	-	53.3	57.5	57.6	53.4	58.1	57.9
Window	-	-	20.3	22.6	25.4	13.0	16.9	19.5
Various	-	-	19.9	20.6	21.0	11.6	12.2	12.1
Average	56.61	-	47.26	49.22	49.84	46.68	48.56	49.04
Overall	67.33	71.28	71.52	72.9	73.46	74.1	74.62	75.23
IoU	-	35.96	37.01	38.69	39.36	37.74	38.96	39.57

Table 4. Segmentation results of various methods on labelmeFacades dataset.

Class	Auto Context (AC)			AC + Potts Model			[5]	[11]
	ST1	ST2	ST3	PW1	PW2	PW3		
Door	64.9	69.4	69.7	65.0	69.6	69.7	59	57
Shop	93.9	95.3	95.4	94.5	95.9	95.9	88	97
Balcony	70.2	76.8	77.7	70.7	77.2	78.0	63	82
Window	75.1	79.9	81.2	75.2	79.8	81.4	66	82
Wall	90.7	91.3	91.2	92.5	92.4	92.3	84	89
Sky	96.5	97.1	97.5	96.9	97.4	97.7	92	98
Roof	74.4	77.9	77.3	76.9	79.4	78.7	58	81
Average	80.83	83.97	84.28	81.67	84.54	84.83	72.9	83.76
Overall	85.88	88.08	88.29	86.96	88.79	89.03	78	88.8
IoU	68.32	72.03	72.39	69.85	73.16	73.51	58	-

Table 5. Segmentation results of various methods on Art-deco dataset.

	2D							3D							2D+3D				
	[8]-1	ST1	ST2	ST3	PW1	PW2	PW3	[6]	[14]-1	ST1	ST2	PW1	PW2	[14]-2	ST3	ST4	PW3	PW4	
Window	-	70.1	68.2	69.0	71.0	69.1	69.0	-	-	58.7	69.6	56.9	67.9	-	69.2	71.7	66.7	72.5	
Wall	-	82.4	83.8	84.6	83.7	85.2	84.6	-	-	83.4	71.6	84.0	73.4	-	87.4	84.3	88.7	85.2	
Balcony	-	80.4	81.0	83.2	80.5	81.0	83.2	-	-	61.7	68.6	61.0	69.8	-	78.9	84.9	76.1	85.1	
Door	-	29.0	32.8	26.0	29.0	33.3	26.0	-	-	0.0	0.0	0.0	0.0	-	25.2	60.1	23.7	61.3	
Roof	-	74.7	73.8	75.7	75.9	76.4	75.7	-	-	75.0	69.6	79.3	71.6	-	78.2	74.0	81.4	76.0	
Sky	-	93.1	93.7	94.5	94.1	94.9	94.5	-	-	95.8	96.5	96.8	97.1	-	94.8	95.6	95.3	96.4	
Shop	-	78.4	76.0	83.1	80.9	78.7	83.1	-	-	100.0	100.0	100.0	100.0	-	85.8	74.5	89.0	76.6	
Average	-	72.59	72.76	72.66	73.57	74.09	73.72	-	-	68.0	67.79	68.29	68.54	-	74.22	77.86	74.40	79.01	
Overall	-	79.12	79.39	79.98	80.37	80.79	81.21	-	-	78.22	81.95	82.30	79.19	-	82.72	80.87	83.35	81.91	
IoU	57.53	58.13	58.37	58.89	59.63	60.18	60.46	42.34	53.22	54.29	56.30	57.04	55.13	61.95	61.95	61.16	62.68	62.64	
Runtime	379	27	56	85	59	88	117	15	21	19	19	19	20	404	85	114	117	146	

Table 6. Segmentation results of various methods for the image labeling task on RueMonge2014 dataset. The runtimes shown here, in minutes, is the time taken to segment the entire dataset and includes the feature extraction, classification and optional projection.

	2D									3D							2D+3D					
	[8]	[14]-1	[14]-2	ST1	ST2	ST3	PW1	PW2	PW3	[6]	[14]-3	[14]-4	ST1	ST2	PW1	PW2	[14]-5	[14]-6	ST3	ST4	PW3	PW4
Window	-	-	-	69.1	72.7	71.2	68.9	72.8	71.1	-	-	-	62.8	68.8	62.7	68.3	-	-	70.9	75.4	69.6	74.5
Wall	-	-	-	83.9	85.3	86.7	85.7	86.3	87.7	-	-	-	91.4	86.2	91.9	87.1	-	-	90.3	89.7	90.8	90.2
Balcony	-	-	-	79.5	78.3	79.2	78.4	77.8	78.5	-	-	-	46.5	56.9	46.9	57.1	-	-	78.5	78.8	78.6	78.9
Door	-	-	-	24.8	22.2	26.3	26.6	21.7	26.5	-	-	-	12.7	32.5	11.5	33.0	-	-	18.8	32.7	19.1	31.2
Roof	-	-	-	65.5	71.0	74.2	66.1	70.1	74.0	-	-	-	70.3	68.7	71.4	70.3	-	-	76.3	75.5	76.9	76.2
Sky	-	-	-	94.5	94.2	94.5	94.5	94.5	94.8	-	-	-	88.0	88.8	88.8	89.0	-	-	95.3	94.7	95.2	94.7
Shop	-	-	-	77.9	79.3	76.2	78.6	79.8	76.4	-	-	-	74.1	74.2	74.7	74.6	-	-	83.1	81.1	83.5	81.5
Average	-	-	-	70.75	71.86	72.60	71.28	71.87	72.72	-	-	-	63.67	68.01	63.97	68.49	-	-	73.33	75.41	73.40	75.33
Overall	-	-	-	79.32	80.91	81.59	80.30	81.41	82.07	-	-	-	78.84	77.94	79.29	78.62	-	-	85.25	84.38	84.49	84.67
IoU	56.10	55.72	55.39	55.66	57.10	58.21	56.68	57.49	58.63	42.32	52.09	52.24	51.34	53.60	51.84	54.36	60.05	60.83	60.59	62.72	60.87	62.88
Runtime	382	302	380	28	57	86	60	89	118	15	15	23	15	15	15	16	317	325	86	86	86	87

Table 7. Segmentation results of various methods for the task of point cloud labeling on RueMonge2014 dataset. The runtimes shown here, in minutes, is the time taken to segment the entire dataset and includes the feature extraction, classification and optional projection.

	2D							2D+3D			
	[6]	ST1	ST2	ST3	PW1	PW2	PW3	ST3	ST4	PW3	PW4
Window	-	72.8	70.9	71.2	72.8	70.8	71.6	72.1	74.5	68.6	74.7
Wall	-	87.1	88.3	87.6	87.7	89.0	88.1	90.9	88.0	91.9	88.5
Balcony	-	76.4	77.2	80.0	75.8	76.4	79.4	75.1	83.2	71.2	82.5
Door	-	19.7	25.2	17.6	19.2	26.5	16.7	18.8	62.9	18.6	62.6
Roof	-	73.8	76.5	75.4	73.0	76.2	75.3	77.7	74.5	79.5	75.0
Sky	-	92.9	93.3	93.3	93.4	93.7	93.8	94.2	95.5	94.2	95.8
Shop	-	78.7	76.2	80.5	79.1	76.3	80.4	83.5	80.8	84.1	81.1
Average	-	71.63	72.51	72.21	71.57	72.72	72.19	73.19	79.91	72.57	80.03
Overall	-	81.73	82.31	82.50	82.00	82.65	82.77	84.44	84.11	84.34	84.42
IoU	41.92	57.81	58.97	58.65	57.94	59.28	58.83	60.88	63.19	60.65	63.66
Runtime	15	31	60	89	63	92	121	89	118	121	150

Table 8. Segmentation results of various methods for the task of mesh labelling on RueMonge2014 dataset. The runtimes shown here, in minutes, is the time taken to segment the entire dataset and includes the feature extraction, classification and projection.

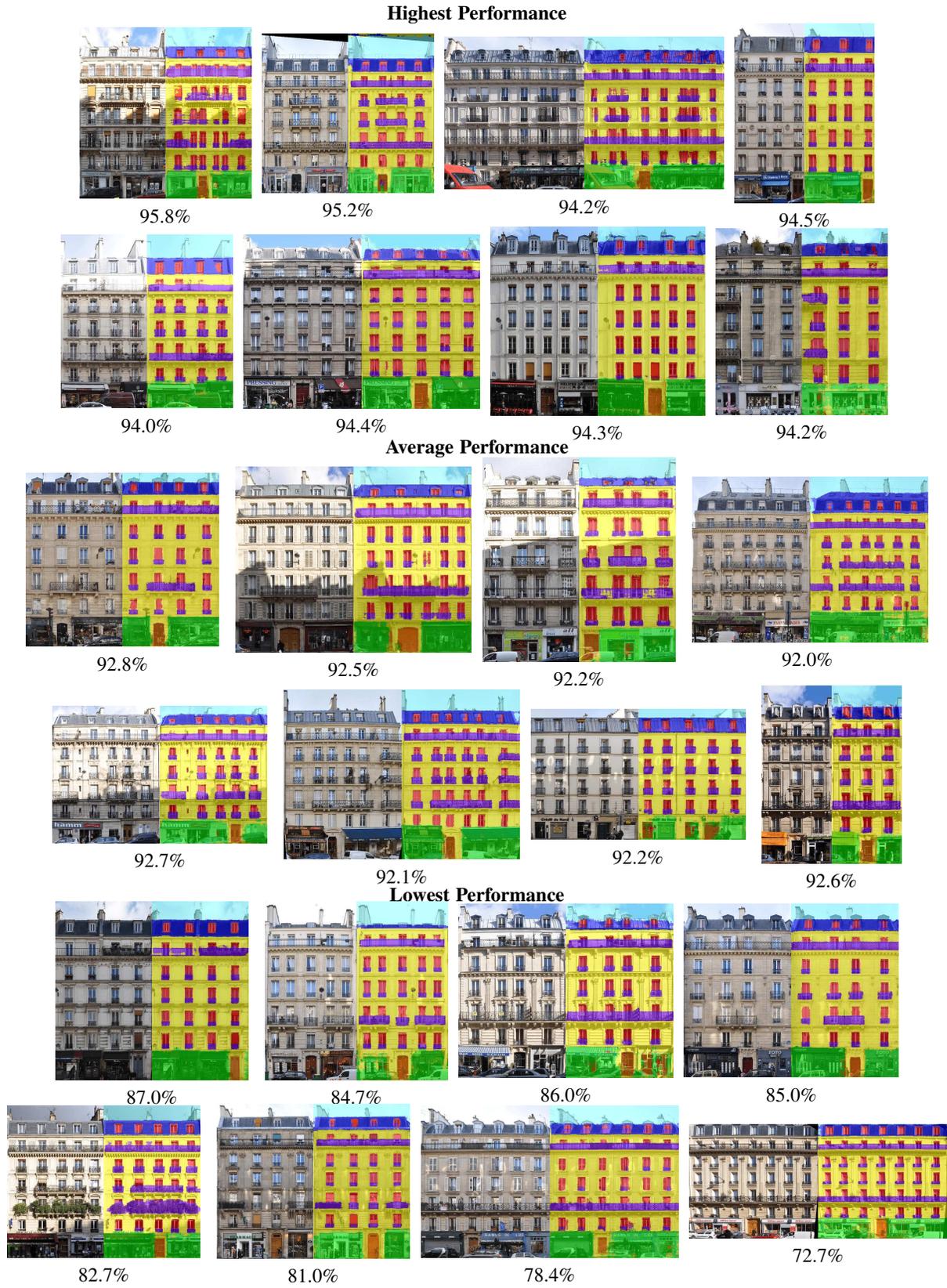


Fig. 2. Qualitative results on ECP dataset images along with overall pixel accuracy (Stage-3 results).



Fig. 3. Qualitative results on Graz dataset images along with overall pixel accuracy (Stage-3 Results).



Fig. 4. Qualitative results on eTRIMS dataset images along with overall pixel accuracy (Stage-3 results).



Fig. 5. Qualitative results on labelmeFacades dataset images along with overall pixel accuracy (Stage-3 results).



Fig. 6. Qualitative results on CMP dataset images along with overall pixel accuracy (Stage-3 results).



Fig. 7. Qualitative results on Artdeco dataset images along with overall pixel accuracy (Stage-3 results).

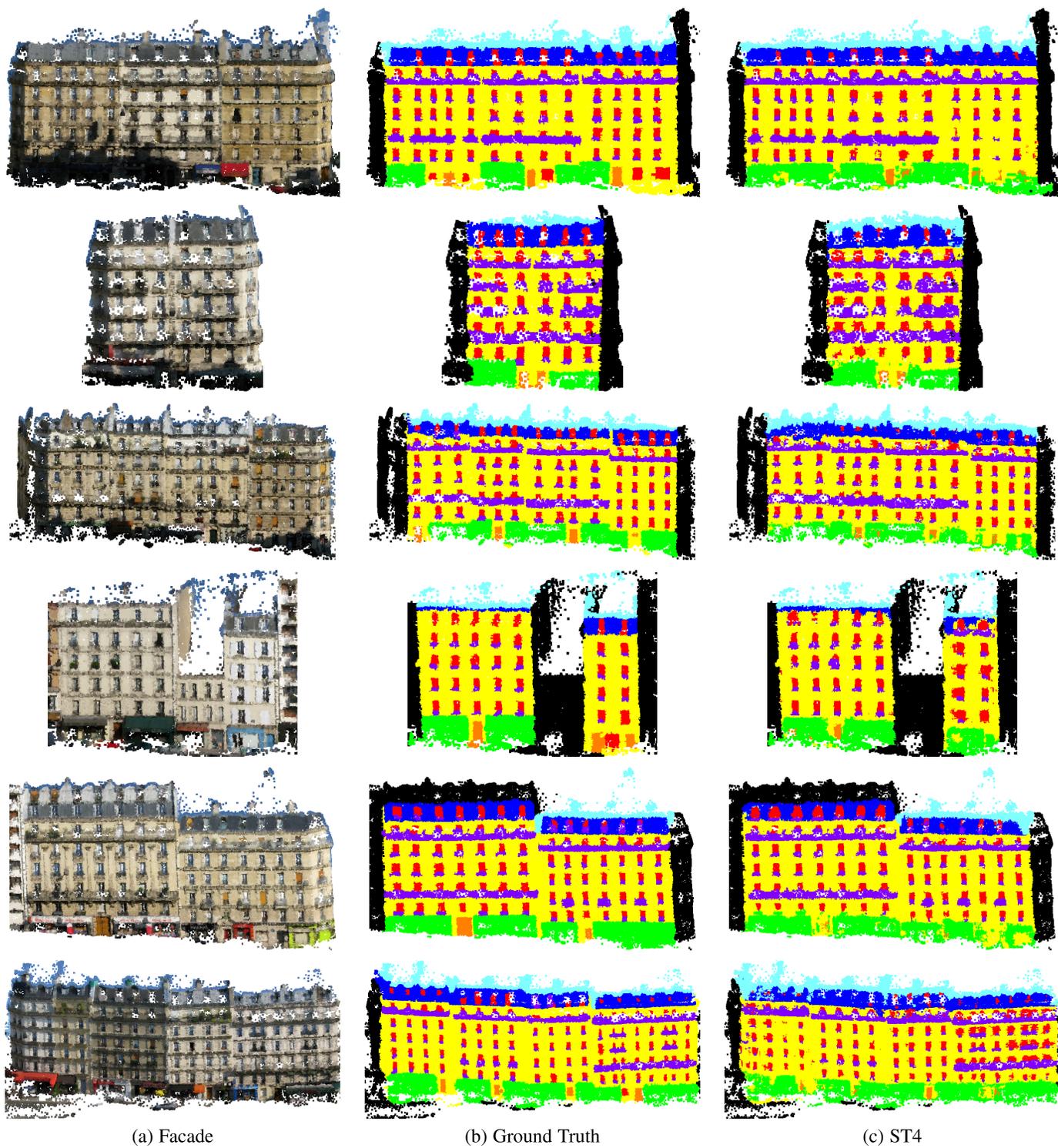


Fig. 8. RueMonge2014 point cloud labeling results.

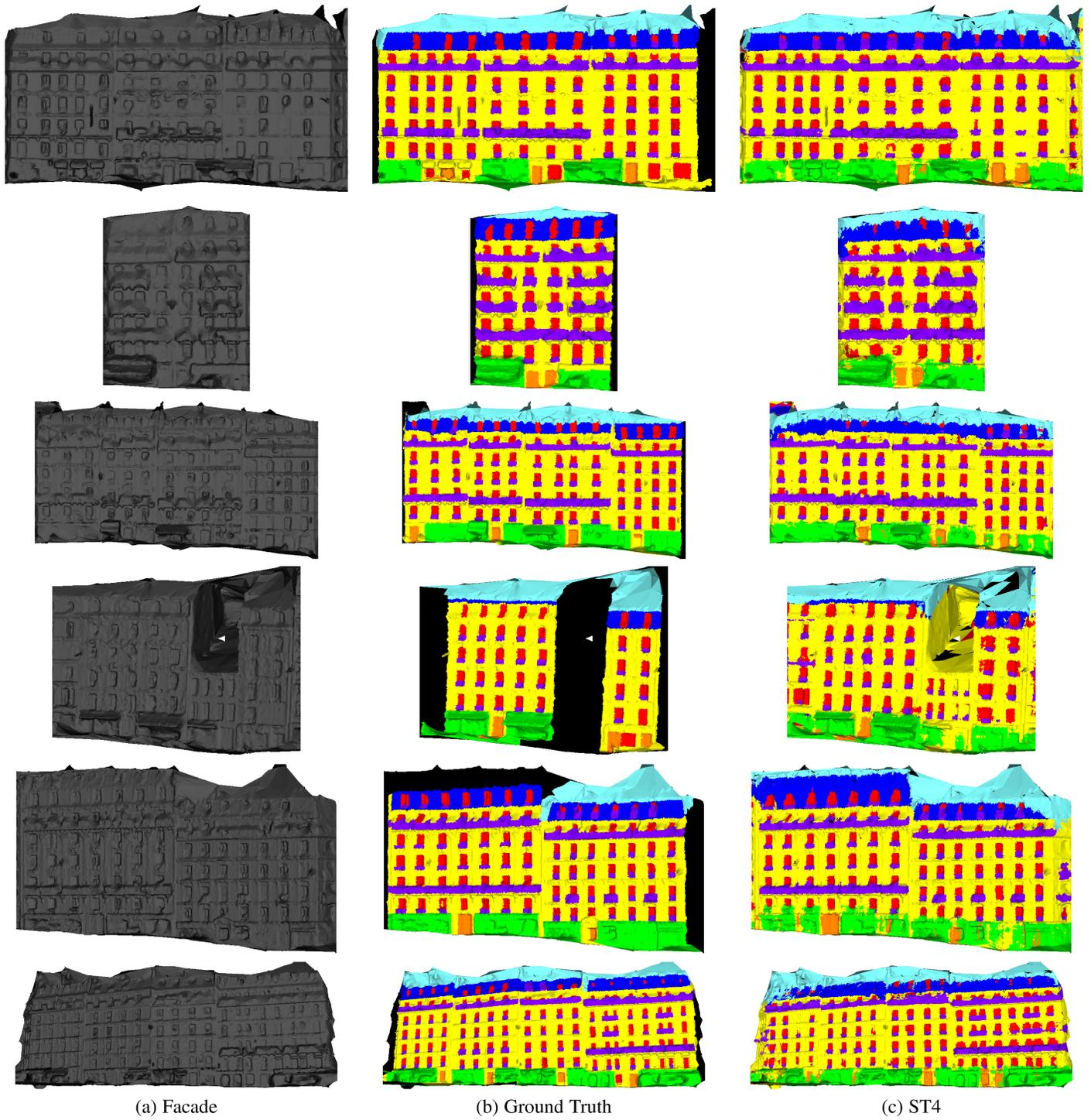


Fig. 9. RueMonge2014 mesh labeling results.

REFERENCES

- [1] F. Korc and W. Forstner, "eTRIMS Image Database for interpreting images of man-made scenes," Dept. of Photogrammetry, University of Bonn, Tech. Rep. TR-IGG-P-2009-01, April 2009.
- [2] R. Tylecek and R. Sara, "Spatial pattern templates for recognition of objects with regular structure," in *GCPR*, 2013.
- [3] H. Riemenschneider, U. Krispel, W. Thaller, M. Donoser, S. Havemann, D. Fellner, and H. Bischof, "Irregular lattices for complex shape grammar facade parsing," in *CVPR*, 2012.
- [4] B. Fröhlich, E. Rodner, and J. Denzler, "A fast approach for pixelwise labeling of facade images," in *ICPR*, 2010.
- [5] R. Gadde, R. Marlet, and N. Paragios, "Learning grammars for architecture-specific facade parsing," *IJCV*, 2016.
- [6] H. Riemenschneider, A. Bódis-Szomorú, J. Weissenberg, and L. Van Gool, "Learning where to classify in multi-view semantic segmentation," in *ECCV*, 2014.
- [7] O. Teboul, I. Kokkinos, L. Simon, P. Koutsourakis, and N. Paragios, "Shape grammar parsing via reinforcement learning," in *CVPR*, 2011.
- [8] A. Martinovic, M. Mathias, J. Weissenberg, and L. Van Gool, "A three-layered approach to facade parsing," in *ECCV*, 2012.
- [9] A. Cohen, A. G. Schwing, and M. Pollefeys, "Efficient structured parsing of facades using dynamic programming," in *CVPR*, 2014.
- [10] C. Gatta and F. Ciompi, "Stacked sequential scale-space taylor context," *T-PAMI*, 2014.
- [11] M. Kozinski, R. Gadde, S. Zagoruyko, R. Marlet, and G. Obozinski, "A MRF shape prior for facade parsing with occlusions," in *CVPR*, 2015.
- [12] B. Fröhlich, E. Rodner, and J. Denzler, "Semantic segmentation with millions of features: Integrating multiple cues in a combined random forest approach," in *ACCV*, 2012.
- [13] S. Nowozin, "Optimal decisions from probabilistic models: the intersection-over-union case," in *CVPR*, 2014.
- [14] A. Martinovic, J. Knopp, H. Riemenschneider, and L. Van Gool, "3d all the way: Semantic segmentation of urban scenes from start to end in 3d," in *CVPR*, 2015.