Fast segmentation of RGB-D images for semantic scene understanding

Dirk Holz, **Alexander J. B. Trevor**, Michael Dixon,
Suat Gedikli, Radu B. Rusu, and Sven Behnke

universität bonn ais
Georgia Institute of Technology
Willow Garage

Problem

As robots move away from pre-programmed action sequences in controlled laboratory setups towards complex tasks in real-world scenarios, both the perception capabilities of these systems and their abilities to acquire and model semantic information must become more powerful. In this context, fast means for pre-processing acquired sensory information and segmenting task-relevant regions are an enabling technology and a prerequisite for avoiding longer delays in sense-plan-act cycles.

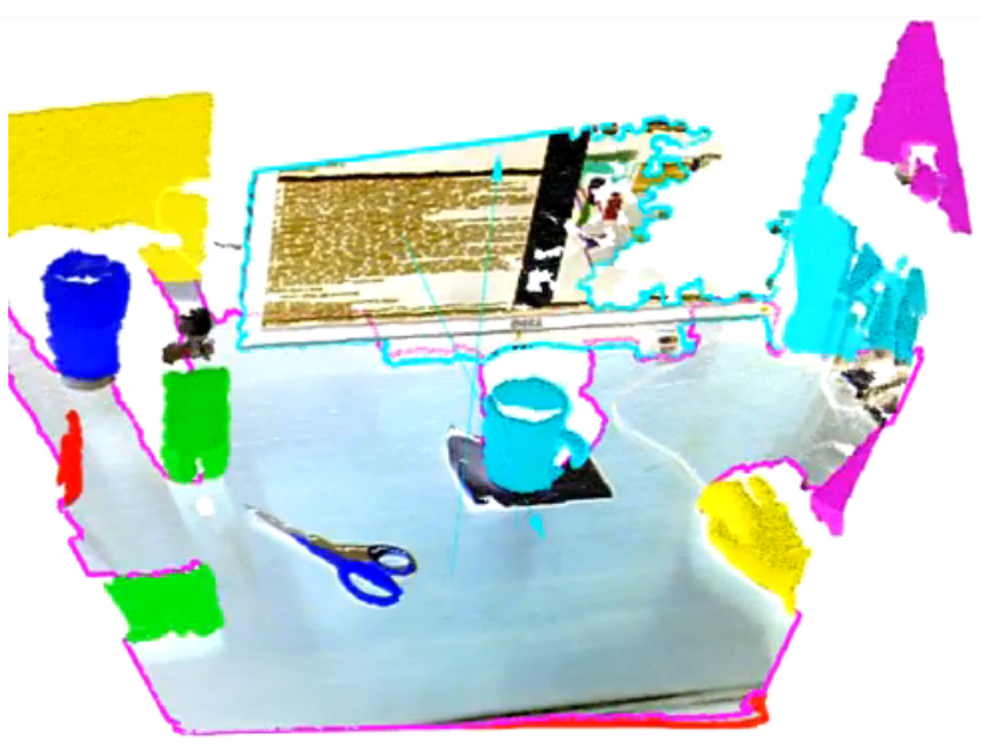
We present two fast segmentation methods for RGB-D images:

1. **Direct image segmentation**
2. Approximate surface reconstruction + **mesh segmentation**

1. Image Segmentation

The first method also exploits the image structure by using a connected component based technique. Each pixel is compared to neighboring pixels (in a 4-connected sense) using a comparison function. Points are considered part of the same segment if the comparison function returns true.

Different comparison functions can be used for different segmentation tasks, such as a plane equation comparison (the dot product between normals and range must match), euclidean distance, color, or combinations of these. These can be run in a sequence and with an optional mask, for example: tabletop objects can be detected by first detecting planar regions, then using these regions as a mask and segmenting with an euclidean distance comparison.



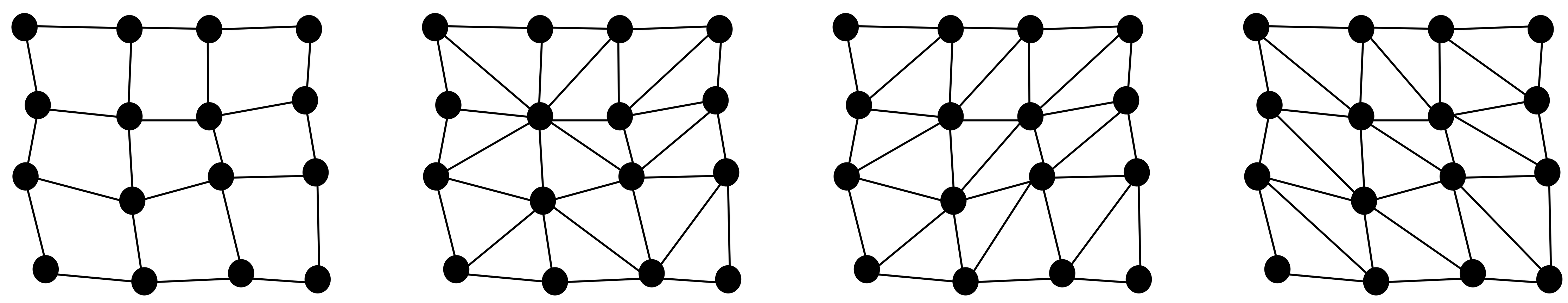
LIVE DEMO!

More Information

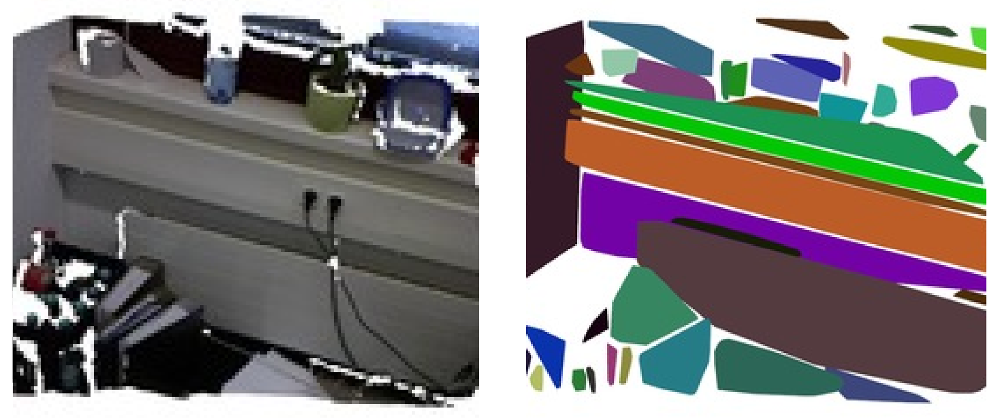
See the **P**oint **C**loud **L**ibrary **PCL** [1] for further details, documentation, and open source implementations:
`http://pointclouds.org`

2. Approx. Surface Reconstruction + Segmentation

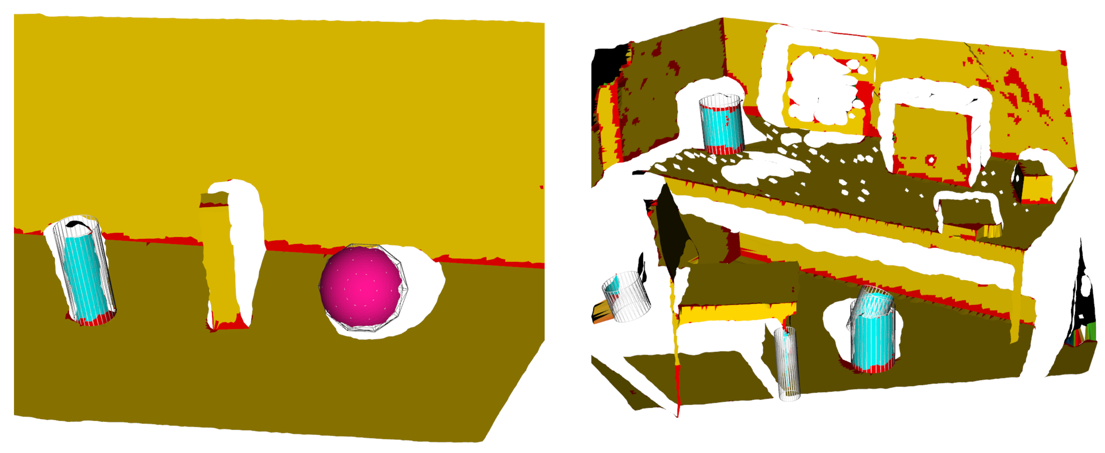
In the second method, we exploit the organized structure of RGB-D images and apply an approximate surface reconstruction [2] by simply connecting adjacent image pixels.



The resulting mesh efficiently caches local neighborhoods for further processing. Furthermore, we compute approximate surface normals directly on the mesh and apply a multi-lateral filtering step to considerably smooth the data. Using an efficient region growing implementation and different region models, we can efficiently compute plane segmentations and full polygonalizations, or segment locally smooth regions and detect geometric shape primitives.



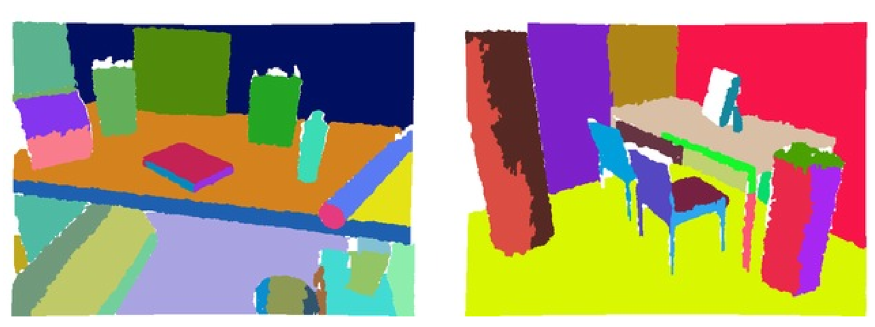
Plane segmentation and polygonalization



Segmentation and primitive detection

SegComp PERCEPTRON data set (30 test images) from [3], (80 % pixel overlap)

Approach	correctly detected	orientation deviation	over-seg.	under-seg.	missed	false / det.
USF [4]	8.9 (60.9%)	2.7	0.4	0.0	5.3	3.6
WSU [4]	5.9 (40.4%)	3.3	0.5	0.6	6.7	4.8
UB [4]	9.6 (65.7%)	3.1	0.6	0.1	4.2	2.8
UE [4]	10.0 (68.4%)	2.6	0.2	0.3	3.8	2.1
UFPR [4]	**11.0 (75.3%)**	**2.5**	**0.3**	**0.1**	**3.0**	**2.5**
Ours	**11.0 (75.3%)**	**2.6**	**0.4**	**0.2**	**2.7**	**0.3**



Fast Plane Segmentation

Resolution	160 × 120	320 × 240	640 × 480
Meshing	11 ms	45 ms	188 ms
Normals	2 ms	7 ms	33 ms
Filtering	10 ms	38 ms	155 ms
Segmentation	6 ms	30 ms	126 ms

Runtimes: 1000 runs on an Intel Core 2 DUO 2.26 GHz (no parallelization)

References

[1] R. B. Rusu and S. Cousins, “3D is here: Point Cloud Library (PCL),” in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, 2011, pp. 1–4.

[2] D. Holz and S. Behnke, “Fast Range Image Segmentation and Smoothing using Approximate Surface Reconstruction and Region Growing,” in *Proceedings of the International Conference on Intelligent Autonomous Systems (IAS)*, Jeju Island, Korea, 2012.

[3] A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Flynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, and R. B. Fisher, “An experimental comparison of range image segmentation algorithms,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 673–689, 1996.

[4] P. Gotardo, O. Bellon, and L. Silva, “Range image segmentation by surface extraction using an improved robust estimator,” in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Madison, WI, USA, 2003, pp. 33–38.