Constrained Planar Cuts - Object Partitioning for Point Clouds

Markus Schoeler, Jeremie Papon and Florentin Wörgötter
III Physikalisches Institut - Biophysik, Georg-August University of Göttingen.

Segmentation of 3D objects into functional parts - forming a visual hierarchy - is a fundamental task in computer vision. Visual hierarchies are essential for many higher level tasks such as activity recognition [5], semantic segmentation [1], object detection [2], and human pose recognition [7]. Nevertheless, part segmentation, particularly of 3D point clouds, remains an open area of research - as demonstrated by the inability of state-of-the-art methods to match human performance on existing benchmarks without excessive fitting to particular ground-truth training examples.

In this work, we aim to partition objects from the bottom-up using a purely geometric approach that generalizes to most object types. This is in stark contrast to recent learning-based methods, which achieve good performance by training separate classifiers for each object class [3, 6]. While such methods do perform well on benchmarks, they are severely restricted in that one must know the object class a-priori, and they do not generalize to new objects at all. With unsupervised methods, such as the one presented in this work, there is no need to create new training data and annotated ground truth, allowing them to be employed as an off-the-shelf first step in object partitioning.

For this we use local concavity information to find semi-global euclidean planar cuts which match a concave boundary model. To find such cuts we propose a directionally weighted, locally constrained sample consensus scheme which is being applied to the edges extracted from the Supervoxel Adjacency Graph [4]. While being robust to noise it uses high weights for concave edges and penalties for convex edges in a locally constrained model evaluation phase, which leads to remarkably accurate partitionings of objects as seen in Fig. 1. We were able to achieve better than state-of-the-art results compared to all published results from unsupervised or weakly-supervised methods and even compete with some data-driven supervised methods. We also introduced a protocol to adapt mesh-segmentation benchmarks to point clouds using an equi-density randomized point sampling, and a back-propagation of found labels to the mesh. This allowed us to report the first quantitative and qualitative results on part-segmentation for point clouds.

Despite being a purely data-driven method our algorithm can cope with shape variations and added Gaussian noise resulting in consistent segmentations. This means that the method is directly applicable as the first step in an automated bootstrapping process and can segment arbitrary unknown objects.

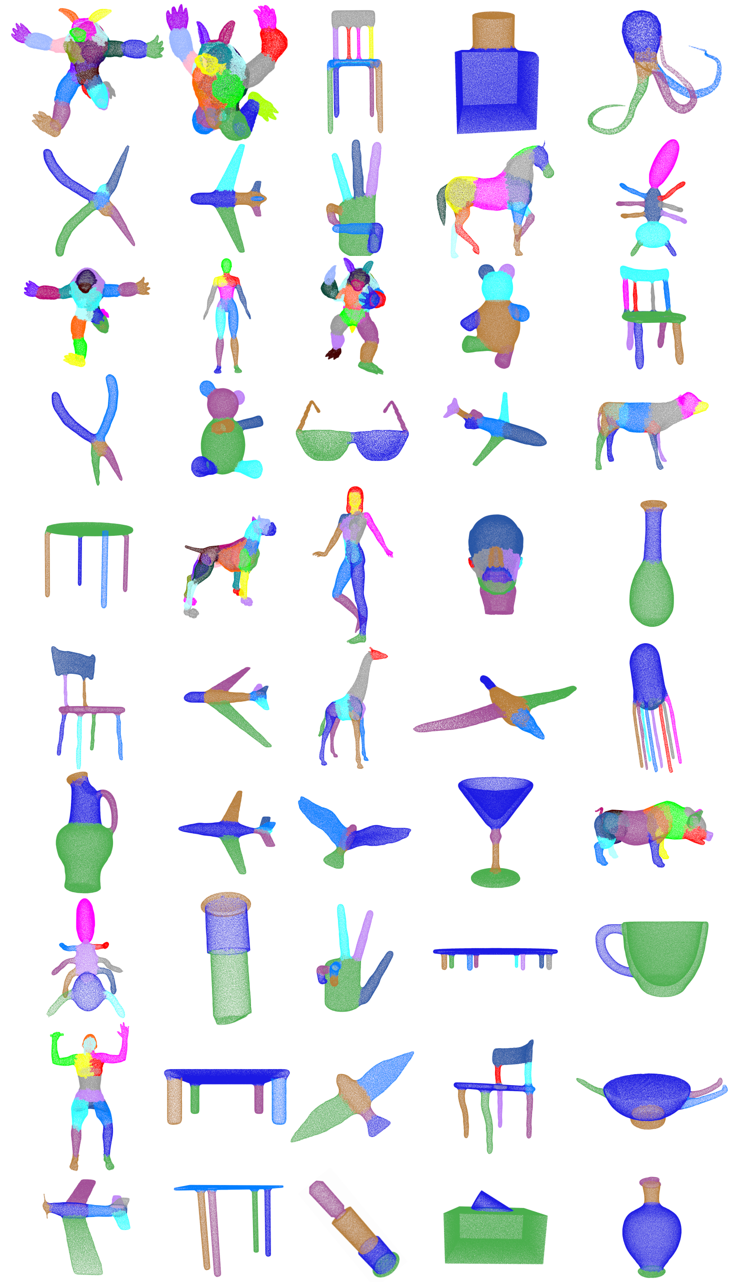

Figure 1: Qualitative results on the Princeton benchmark using fixed parameters.

[1] P. Arbelaez, B. Hariharan, Chunhui Gu, S. Gupta, L. Bourdev, and J. Malik. Semantic segmentation using regions and parts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3378–3385, 2012.

[2] Ali Farhadi and Mohammad Amin Sadeghi. Phrasal recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12): 2854–2865, December 2013.

[3] Evangelos Kalogerakis, Aaron Hertzmann, and Karan Singh. Learning 3d mesh segmentation and labeling. *ACM Transactions on Graphics*, 29(4):1, July 2010.

[4] J. Papon, A. Abramov, M. Schoeler, and F. Wörgötter. Voxel cloud connectivity segmentation - supervoxels for point clouds. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[5] Hamed Pirsiavash and Deva Ramanan. Detecting activities of daily living in first-person camera views. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2847–2854. IEEE, 2012.

[6] Weiwei Xu, Zhouxu Shi, Mingliang Xu, Kun Zhou, Jingdong Wang, Bin Zhou, Jinrong Wang, and Zhenming Yuan. Transductive 3d shape segmentation using sparse reconstruction. *Computer Graphics Forum*, 33(5):107–115, August 2014.

[7] Yi Yang and Deva Ramanan. Articulated pose estimation with flexible mixtures-of-parts. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1385–1392, 2011.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.