Semantic Segmentation of Outdoor Areas using 3D Moment Invariants and Contextual Cues

Sven Sickert[1], and Joachim Denzler[1,2]

[1]Computer Vision Group, Friedrich Schiller University Jena, Germany
[2]Michael Stifel Center Jena, Germany

Abstract. In this paper, we propose an approach for the semantic segmentation of a 3D point cloud using local 3D moment invariants and the integration of contextual information. Specifically, we focus on the task of analyzing forestal and urban areas which were recorded by terrestrial LiDAR scanners. We demonstrate how 3D moment invariants can be leveraged as local features and that they are on a par with established descriptors. Furthermore, we show how an iterative learning scheme can increase the overall quality by taking neighborhood relationships between classes into account. Our experiments show that the approach achieves very good results for a variety of tasks including both binary and multi-class settings.

1 Introduction

In recent years, several new techniques for the creation of 3D point clouds have emerged. Current structure from motion techniques like LSD-SLAM [2] raise each webcam to a powerful 3D scanning device. Furthermore, in companies and research groups the use of laser based devices like LiDAR scanners (*light detection and ranging*) becomes more common. Today, this recording technique is also advanced enough to allow for a fast and easy scanning procedure. Thus, the demand increases for algorithms which are able to process such data with respect to typical pattern recognition tasks like detection, segmentation or classification.

In biological research projects like [19,6] with a focus on the ecological system of forests such scanners are used to create a 3D representation of an area. It opens up the possibility to use computer-assisted systems for their analysis. However, there is often the need for a certain amount of user interaction. While this might be feasible in environmental research, other areas like autonomous driving or robotics in general are in need of real-time and thereby fully-automatic systems.

In this paper, we provide an approach for the fully-automatic semantic segmentation of 3D point cloud data recorded by LiDAR scanning devices. We make use of 3D moment invariants [9] which are a powerful local representation and are invariant to Euclidean and affine transformations. For regular 2D image data semantic segmentation represents the pixel-by-pixel classification of a whole scene leading to a segmentation into meaningful regions [3,11,15,16]. The classification is based on local pixel representations within a certain neighborhood.

(a) (b)

Fig. 1. Datasets for the analysis of outdoor areas used in our experiments: (a) 3DForest dataset [19] with labels for *tree*, *terrain*, *dead wood* and *miscellaneous* and (b) Oakland 3D Point Cloud Dataset [12] with labels for *facade*, *ground*, *vegetation*, *wire* and *pole*.

While in grid data this neighborhood can easily be retrieved using common faces or corners of pixels the neighborhood in an orderless point cloud is not as trivial.

We show how to create a local feature descriptor based on moment invariants and augmented contextual cues that can represent neighborhood relations of present classes. For the latter we make use of a cascade of classifiers and the concept of *auto-context* [11,16,21]. Our experiments demonstrate the power of our approach for the automatic segmentation of individual trees, as well as the analysis of whole outdoor areas in general.

The outline of this paper is as follows. We first give a short overview of related work in this field of research. In Section 3 we describe how local 3D moment invariants as a feature descriptor can be derived. The iterative classification scheme and the use of context information during learning is described in Section 4. An evaluation of the proposed framework follows in Section 5. A summary in Section 6 concludes the paper.

2 Related Work

Segmentation tools provided with [6] and [19] for the analysis of forestal areas are based on local statistics and heuristics. Especially, the clustering routines of [6] are mainly based on outlier estimation and distance thresholding. These are basic features for tasks like unsupervised segmentation that are not invariant with respect to scaling and change of density in the data. An analysis of a variety of such local statistics and features can be found in [22].

In a more recent approach local concavity is used as an indicator for boundaries to achieve a bottom-up segmentation [14]. Supervoxels have to be generated in a first step. The authors of [5] propose a contour detection method to find regions of interest. In both [14] and [5] graph-cut techniques are applied to finally segment the point cloud based on the features. In our work, however, we also try to find the meaning of these segmented regions which places additional demands on the features itself.

Many works of the past decade focused on the development of meaningful 3D feature descriptors for classification tasks. In [4] and [17] neighborhood spheres are defined to create histograms of local features. For the 3D shape context descriptor [4] the distribution of points within individual parts of the sphere are used. The histograms of the SHOT descriptor [17] represent the distribution of normal vectors in these parts. A similar work is based on spinning planes around a point along all three axes in order to get a distribution of the residuals [7,8]. In all cases the main direction of the descriptor has to be computed in order to make them invariant with respect to rotations. We propose to use features that are derived for the purpose of being invariant to Euclidean and affine transformations.

For a task like semantic segmentation the modeling of context information is possible. Relationships between nearby points can be directly modeled using graphical models like *Markov random fields* (MRF) [10,13] or *Markov networks* [12]. In the latter, a functional gradient approach is used to increase the performance of MRFs by learning high-order interactions. Additionally, in [18] it was shown how a classification approach based on *conditional random fields* can be extended to an online learning setting that is able to improve iteratively.

Such a system can also be used to integrate classification outputs of previous predictions. In [23] a sequence of classifiers is trained to iteratively refine the learned model by adding context information. The authors of [3] propose to train *random decision forests* (RDF) in an incremental manner using contextual cues in deeper levels to semantically segment single images and image stacks. We adopt this concept for the processing of 3D point clouds by augmenting our initial set of features with information on the class distributions in a certain neighborhood. This allows for modeling of relationships between both points and classes. We also use RDFs in our framework as they are fast to learn and only have very few setup parameters.

3 Local Features Based on 3D Moment Invariants

In order to classify each individual point, we need a powerful representation for it which is especially robust against scaling as well as rotational and translational transformations in 3D space. A natural choice are statistical moments and their associated invariants. In this section, we describe how local 3D moment invariants can be derived for our purpose. As a prerequisite, we define a set $\mathcal{P} \subset \mathbb{R}^3$ of 3D points $\mathbf{p}^{(i)}$, with $1 \leq i \leq N$ and $N = |\mathcal{P}|$ the amount of points in the cloud.

3.1 3D Surface Moments

Let us assume there is a surface triangulation $\mathcal{S}$ given representing the object or scenery using $\mathcal{P}$. Furthermore, $\mathcal{S}$ consists of triangles $\mathcal{T}^{(j)}$, with $1 \leq j \leq N_{\mathcal{T}}$

and $N_{\mathcal{T}} = |\mathcal{S}|$ being the amount of triangles. Such a surface representation can be efficiently created using a *Delaunay* triangulation.

Each triangle consists of three corner points $\mathbf{c}_1^{(j)}, \mathbf{c}_2^{(j)}, \mathbf{c}_3^{(j)} \in \mathcal{P}$. The 3D surface moments M_{kln} of $(k+l+n)^{\text{th}}$ order for $\mathcal{S}$ are defined as the accumulated surface moments of the associated triangles $\mathcal{T}^{(j)}$: $M_{kln} = \sum_j m_{kln}^{(j)}$. In the following we skip the superscript for easier readability. The surface moment m_{kln} for a triangle $\mathcal{T}$ can be computed using

$$m_{kln} = \int\int_{\mathcal{T}} x^k y^l z^n \rho(x,y,z) \mathrm{d}s \quad , \tag{1}$$

where ρ is a density function, with $\rho(x,y,z) = 1$ in our case.

As was shown in [20,24] the calculation of m_{kln} can be reduced to the computation of the area moments

$$m_{pq} = \int\int_{D} u^p v^q \mathrm{d}u \mathrm{d}v \quad , \tag{2}$$

where $u, v \in D \subset \mathbb{R}^2$ and $P_{\mathcal{T}}(u,v)(x_{\mathcal{T}}, y_{\mathcal{T}}, z_{\mathcal{T}})$ a suitable parametrization. For details on the parametrization and a derivation on how to exactly compute m_{kln} using m_{pq} we refer to [20].

3.2 Local 3D Moment Invariants

By calculating M_{kln} using the accumulated surface moments $m_{kln}^{(i)}$ we are now able to compute the eleven 3D moment invariants $I_{22}^2, I_{222}^2, \ldots, I_{1113}^3$ which were originally proposed by Lo & Don in [9]. The authors present moment invariants of second and third order surface moments. Details about their derivation can be found in [9].

In general, these 3D moment invariants can be computed using $\mathcal{P}$ to build a descriptor that characterizes it. However, for a task like semantic segmentation we are interested in a powerful feature representation of each individual point. Hence, we need to compute moment invariants *locally*.

While it would be possible to classify complete objects that are part of a scenery using the moment invariant representation, this would require a very good pre-clustering of $\mathcal{P}$ into objects. However, we can not rely on such a method to be given. Therefor, we propose to represent each point by its local surrounding surface shape. We follow [20] by defining a sphere $S_1^{(i)}$ of radius r_1 around each $\mathbf{p}^{(i)}$. For each $S_1^{(i)}$ we compute the 3D surface and consequently the 3D surface moments using only the 3D points within that sphere. The individual local 3D surface moments can be used to compute individual local 3D moment invariants. We will denote these as features which are part of the vector $\mathbf{x}^{(i)}$.

4 Leveraging Context Information using Random Forests

In this section, we focus on how to augment the features described before with context information using a cascade of classifiers. First, we give a short description of random forests. After that, we explain how intermediate classification outputs can be used to enforce local smoothness and model class relations.

4.1 Random Forests for Semantic Segmentation

Random decision forest (RDF) is a well known machine learning tool that is based on an ensemble of decision trees. Individual decision trees are of limited discriminative power and are also prone to over-fitting during training. Breiman proposed in [1] randomization techniques that can help to overcome several shortcomings of single decision trees. Multiple decision trees are learned using different parts of the whole training data. Additionally, a random sampling of features for the splitting decisions and a subsequent evaluation of these splits increase diversity between the individually learned trees. A final voting among the learned trees yields the final classification result.

In our case each point $\mathbf{p}^{(i)}$ is represented by its local moment invariants in $\mathbf{x}^{(i)} \in \mathbb{R}^{11}$. Accordingly, the splits are based on these feature dimensions only. Local moment invariants are based on a certain neighborhood (see Sec. 3.2) and should thereby be similar for neighboring points. However, in RDFs each example is classified individually without taking classification results of neighboring examples into account. Hence, uncertain areas in the feature space can lead to a partially scattered classification output. In the following we propose an iterative classification scheme that helps to overcome this issue.

4.2 Contextual Cues from Local Neighborhoods

The standard framework of feature extraction and classification provides us with a class decision for each point of the whole point cloud. While this classification result might not be consistent in every detail, it allows to deduce the principal semantics. Otherwise the originally chosen features must have been insufficient for the given task or the classifier was configured poorly. Inspired by [3] we propose to use these results to augment the original feature vector with local class distributions within a certain neighborhood.

Let $\mathbf{p}^{(i)}$ and $S_1^{(i)}$ be a point and its surrounding sphere as described in Section 3.2. Furthermore, let $\mathcal{P}^{(i)} \subset \mathcal{P}$ be the set of points within $S_1^{(i)}$. The amount of examples classified as class c is represented by $N_c^{(i)}$, with $1 \leq c \leq C$ and C the amount of classes. Hence, we can retrieve the relative frequencies for class c by $f_c^{(i)} = \frac{N_c^{(i)}}{|\mathcal{P}^{(i)}|}$ and use them as additional features.

In the same manner, we define a second sphere S_2 with radius $r_2 < r_1$ around that point and retrieve the relative class frequencies therein as well. The idea is, to distinguish between the class distribution next to the point from the distribution in an extended neighborhood. While the inner sphere S_2 should enforce *smoothness* among nearby points, the outer sphere S_1 allows for the modeling of *relationships* between occurring classes.

After computing these features we augment the initial feature vector $\mathbf{x}^{(i)}$ with them. In order to avoid concatenation and thereby reallocation of memory we suggest to initialize the $2C$ dimensions as $\frac{1}{C}$. Thus, the dimension $d = 2C+11$ of a feature vector is constant during the whole classification process.

4.3 Cascaded Random Forests

After the feature augmentation step we train a new RDF with the same configuration as before. Hence, the root nodes include all training examples and entirely new trees are built in the process. A continuous learning scheme like in [3] could also be applied to build one entire RDF in a level-wise manner. However, we find it important to have meaningful contextual cues as possible splitting features for all examples early in the training procedure.

In general, the concept of iteratively learning classifiers based on previous outputs is referred to as auto-context [11,16,21]. Such a step-wise learning procedure can be applied multiple times. The additional information input should increase the performance of the classifier. Thus, its output can again be used to refine the features as were described in Sec. 4.2. The performance gain after each iteration is likely to decrease over time. At some point the overall performance might even get worse because of over-fitting. We will show in our experiments that multiple iterations are beneficial in certain scenarios.

5 Experiments

We evaluate our proposed framework using three datasets. First, we apply it to the binary segmentation task *tree* against *background*. After that, we use a public dataset of a LiDAR scanned forestal area to demonstrate the power of 3D moment invariants as features in comparison with other 3D feature descriptors. Furthermore, we analyze the impact of our proposed iterative classification scheme. In the third experiment we show that our method is also applicable to other outdoor scenes like urban areas.

For the evaluation we use the evaluation metrics *precision* and *recall*. These measures account for unbalanced testing datasets. Additionally, we report the *f1 score* which is the harmonic mean of both measures. In multi-class settings we report metrics that are averaged over all classes. In all our experiments we set the radius of the inner sphere as $r_2 = \frac{1}{2}r_1$. The size of r_1 was analyzed during our experiments and is provided in the evaluation sections, respectively. The RDF consists of 20 trees with a minimum amount of 15 examples in each leaf node. All reported performance results are averaged over five runs.

5.1 Segmentation of Individual Trees

Experimental Setup In this series of experiments we want to show how our proposed framework performs for the task of segmenting an individual tree in a noisy recording. We use a point cloud consisting of 180,304 points showing one tree and its surrounding. A visualization of the point cloud can be found in Fig. 2b. The lower half of the scene is heavily distorted by noise. A ground-truth

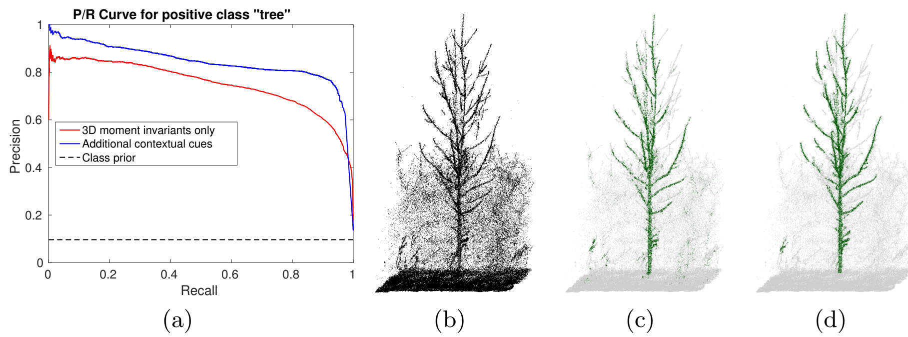


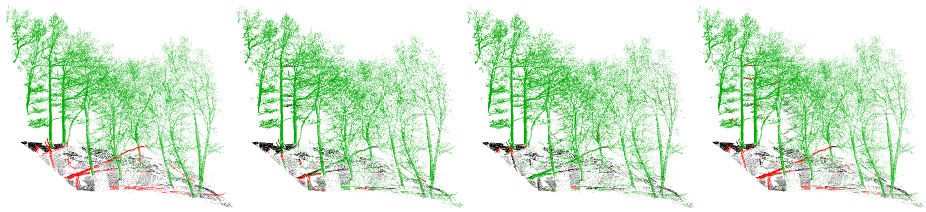
Fig. 2. Results for the segmentation of an individual tree: (a) Precision-recall curve for the testing data. Qualitative results for the whole tree (b) including both training and testing data: (c) only 3D moment invariants and (d) with additional contextual cues.

labeling for the classes *tree* and *background* is available. We split this cloud into two parts in a way that the subsets contain both classes allowing us to report quantitative results. The training set contains 94,488 points and the testing set the remaining 85,816 points.

Evaluation First, we evaluated how 3D moment invariants alone are suited for the segmentation of trees. As can be seen from the precision-recall curve in Fig. 2a the features work already very well for this task. The visualization of the qualitative result in Fig. 2c amplifies this observation. However, in details the results are scattered as was expected (see Sec. 4.1) given the individual classification of each point. By augmenting the raw features with contextual cues we are able to increase the performance. The modeled influence from nearby points and classes has an effect that is both visible in quality and measurable by performance criteria. Especially, the precision increases considerably given our proposed contextual cues.

5.2 Analysis of Forestal Areas

Experimental Setup To evaluate our method on a larger area with multiple trees we use the data of the 3DForest project [19]. It consists of 467,211 points which were recorded using a terrestrial LiDAR scanner. The authors provide labels for the classes *tree*, *terrain* and *dead wood*. Additionally, a background class *miscellaneous* is available. For our experiments we split the data into two parts of almost the same size using the center position along the longest dimension. Points lower and equal to $y = 528.0$ are used for training and the remaining points are used for testing. The complete scenery is depicted in Fig. 1a. Examples of the background class ($\approx 10,000$ points) are excluded from the evaluation.



(a) Ground-truth (b) HOR [7] (c) SHOT [17] (d) Ours

Fig. 3. Qualitative results for 3DForest [19] using different descriptors to classify **tree**, **dead wood** and **terrain**. This figure is best viewed in color (zoom in for details).

	Precision			**Recall**			**F1 Score**
	tree	terrain	dead w.	tree	terrain	dead w.	average
Without context (it. 1)							
HOR [7] ($k = 100$)	0.903	0.676	0.523	0.969	0.664	0.090	0.586
SHOT [17] ($r_1 = 1.7m$)	0.912	0.795	**0.735**	**0.998**	0.721	0.052	0.602
Ours ($r_1 = 0.4m$)	**0.920**	**0.831**	0.551	0.977	**0.723**	**0.276**	**0.696**
Context-based (it. 2)							
HOR [7] ($k = 100$)	0.950	0.663	**0.818**	0.979	0.812	0.226	0.693
SHOT [17]	0.933	0.746	0.654	0.984	**0.873**	0.174	0.665
Ours	**0.960**	**0.862**	0.660	**0.987**	0.797	**0.531**	**0.796**

Table 1. Comparison of different feature descriptors for 3DForest [19] using our proposed framework. For the context-based features a radius $r_1 = 0.5m$ was used.

Evaluation In a first series of experiments we want to know how the 3D moment invariants perform in comparison with established feature descriptors like HOR [7] and SHOT [17]. The pure performance of the features can be seen in the upper part of Tab. 1. For all features the best performing configuration with respect to the neighborhood parameters k and r_1 was used. As can be seen from the results, the SHOT descriptor reaches the best average precision over all classes. In contrast, the average recall using moment invariants is considerably better. In total the averaged f1 score over all classes is best for our proposed descriptor.

Adding the contextual cues improves the performance of all descriptors. However, the advantage of combining moment invariants and context information is obvious. Especially, the class *dead wood* with fewer examples is captured better using our method leading to a higher recall in general. This is also visible from the qualitative results which can be found in Fig. 3. All features can be used to differentiate between *terrain* and *tree*. However, the use of moment invariants performs best on average over all classes.

For a more thorough analysis of our framework we continued with experiments with respect to different parameter settings. The results in Fig. 4a demon-

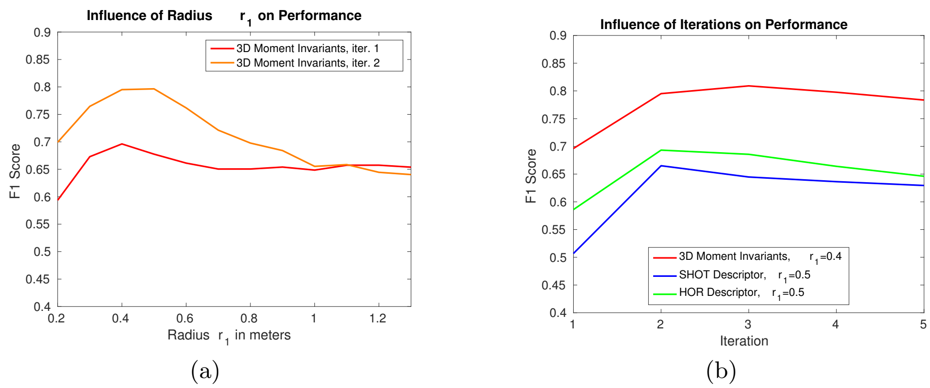


Fig. 4. Quantitative analysis of different feature configurations for 3DForest [19]: (a) Influence of radius r_1 on the performance of the 3D moment invariants and (b) performance depending on the amount of iterations of the cascaded RDF.

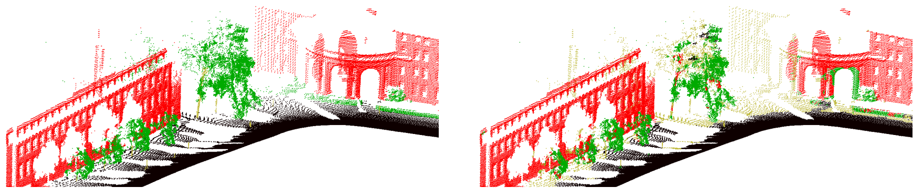
strate how the size of the neighborhood influences the performance of moment invariants in general. Although, the optimal value for parameter r_1 is dependent on the data itself, we observed that a value of 0.5 is very good starting point. A larger value might be necessary for the modeling of class relationships.

In our last series of experiments we look into the iterative classification itself. As can be seen from the plot in Fig. 4b more than two iterations are in most cases not beneficial. However, this is not true for the use of moment invariants as features. The performance improved after the third iteration showing again how well the combination of moment invariants and contextual cues work.

5.3 Urban Scenes

Experimental Setup In our last series of experiments we test our method in a different setting to show its wide applicability. We use the *Oakland 3D Point Cloud Dataset* [12] which contains scenes from an urban area (see Fig. 1b). For the experimental setup we follow the training and testing splits provided in [12]. Thus, we train our approach using 36,932 points with labels of the classes *facade*, *ground*, *pole/trunk*, *wire* and *vegetation*.

Evaluation In contrast to the analysis of forestal areas the task of urban scene understanding contains even more classes with few examples and tiny details. Hence, the modeling of contextual information is even more important. However, the performance of moment invariants as a feature descriptor alone is still interesting. In comparison with other local features without context modeling we are able to outperform the best setup of [22] in terms of precision by almost 7 percentage points. Our performance with respect to recall is only slightly worse. An overview over all results can be found in Tab. 2. We are also able to compete



(a) Ground-truth (b) Ours

Fig. 5. Qualitative result for a partial view of the Oakland 3D Point Cloud Dataset [12] with visible classes **facade**, **vegetation**, **pole** and **ground**. Areas with low spatial density tend to be confused with the class pole because of their structural appearance.

average	Without context [22]	Without context **Ours**	Context-based [13]	Context-based [10]	Context-based [12]	Context-based **Ours**
Precision	0.611	**0.678**	0.566	0.704	0.730	**0.736**
Recall	**0.739**	0.710	0.807	0.866	**0.902**	0.798
F1 Score	0.623	**0.655**	0.587	0.757	**0.778**	0.695

Table 2. Quantitative results for the Oakland 3D Point Cloud Dataset [12]: the measures are averaged over all five classes. Misses for the subtle class wire have a negative effect on our overall recall in comparison with MRF-based approaches.

with state-of-the-art results which include context modeling. The recall for all classes on average is worse than the MRF methods because of false negatives for the class wire. Especially [12], with its highly optimized learning procedure is performing better with repsect to subtle structures. However, in terms of precision and overall performance our proposed method is very well suited for this task. A comparison of our result with the ground-truth can be found in Fig. 5.

6 Conclusions

In this paper, we showed how 3D moment invariants and contextual cues can be combined for the semantic segmentation of outdoor areas. With this approach 3D point clouds created by terrestrial LiDAR scanners can be analyzed in a fully automatic manner. We proposed an iterative classification framework based on powerful local feature descriptors that are invariant to many transformations in 3D space. Furthermore, we were able to overcome the drawback of RDFs for the task of semantic segmentation which often leads to scattered classification results due to its individual classification scheme. Experiments show its power for tasks like the segmentation of individual trees and the analysis of whole forestal or urban areas with multiple classes.

References

1. Breiman, L.: Random forests. Machine Learning 45(1), 5–32 (October 2001)
2. Engel, J., Schöps, T., Cremers, D.: Lsd-slam: Large-scale direct monocular slam. In: Computer Vision - ECCV 2014. Lecture Notes in Computer Science, vol. 8690, pp. 834–849 (2014)
3. Fröhlich, B., Rodner, E., Denzler, J.: Semantic segmentation with millions of features: Integrating multiple cues in a combined random forest approach. In: Asian Conference on Computer Vision. pp. 218–231 (2012)
4. Frome, A., Huber, D., Kolluri, R., Blow, T., Malik, J.: Recognizing objects in range data using regional point descriptors. In: Computer Vision - ECCV 2004. pp. 224–237. Lecture Notes in Computer Science (2004)
5. Hackel, T., Wegner, J.D., Schindler, K.: Contour detection in unstructured 3d point clouds. In: IEEE International Conference on Computer Vision and Pattern Recognition. pp. 1610–1618 (2016)
6. Hackenberg, J., Spiecker, H., Claders, K., Disney, M., Raumonen, P.: Simpletree - an efficient open source tool to build tree models from tls clouds. Forests 6(11), 4245–4294 (November 2015)
7. Krückhans, M.: Ein Detektor für Ornamente auf Gebäudefassaden auf Basis des histogram-of-oriented-gradients- Operators. Master's thesis, Rheinische Friedrich-Wilhelms-Universitt Bonn (2010)
8. Lang, D., Friedmann, S., Paulus, D.: Semantic 3d octree maps based on conditional random fields. In: IAPR International Conference on Machine Vision Applications. pp. 185–188 (2013)
9. Lo, C.H., Don, H.S.: 3-d moment forms: Their construction and application to object identification and positioning. IEEE Transactions on Pattern Analysis and Machine Intelligence 11(10), 1053–1064 (October 1989)
10. Lu, Y., Rasmussen, C.: Simplified markov random fields for efficient semantic labeling of 3d point clouds. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 2690–2697 (2012)
11. Montillo, A., Shotton, J., Winn, J., Iglesias, J.E., Metaxas, D., Criminisi, A.: Entangled decision forests and their application for semantic segmentation of ct images. In: International Conference on Information Processing in Medical Imaging. pp. 184–196 (2011)
12. Munoz, D., Bagnell, J.A., Vandapel, N., Hebert, M.: Contextual classification with functional max-margin markov networks. In: IEEE International Conference on Computer Vision and Pattern Recognition. pp. 975–982 (2009)
13. Munoz, D., Vandapel, N., Hebert, M.: Onboard contextual classification of 3-d point clouds with learned high-order markov random fields. In: IEEE International Conference on Robotics and Automation. pp. 2009–2016 (2009)
14. Schoeler, M., Papon, J., Wörgötter, F.: Constrained planar cuts - object partitioning for point clouds. In: IEEE International Conference on Computer Vision and Pattern Recognition. pp. 5207–5215 (2015)
15. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. In: IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 39, pp. 640–651 (April 2017)
16. Shotton, J., Johnson, M., , Cipolla, R.: Semantic texton forests for image categorization and segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8 (2008)

17. Tombari, F., Salti, S., Stefano, L.D.: Unique signature of histograms for local surface description. In: Computer Vision - ECCV 2010. pp. 356–369. Lecture Notes in Computer Science (2010)
18. Tombari, F., Stefano, L.D., Giardino, S.: Online learning for automatic segmentaton of 3d data. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 4857–4864 (2011)
19. Trochta, J., , Krůček, M., Král, K.: 3d forest. `http://www.3dforest.eu/`, accessed: 2017-03-07
20. Trummer, M., Süße, H., Denzler, J.: Coarse registration of 3d surface triangulations based on moment invariants with applications to object alignment and identification. In: IEEE International Conference on Computer Vision. pp. 1273–1279 (2009)
21. Tu, Z., Bai, X.: Auto-context and its application to high-level vision tasks and 3d brain image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(10), 1744–1757 (October 2010)
22. Weinmann, M., Jutzi, B., Mallet, C.: Feature relevance assessment for the semantic interpretation of 3d point cloud data. In: ISPRS Annals of Photgrammetry, Remote Sensing and Spatial Information Science. vol. II-5/W2, pp. 313–318 (2013)
23. Xiong, X., Munoz, D., Bagnell, J.A., Hebert, M.: 3-d scene analysis via sequenced prediction over points and regions. In: IEEE International Conference on Robotics and Automation. pp. 2609–2616 (2011)
24. Xu, D., Li, H.: 3-d surface moment invariants. In: International Conference on Pattern Recognition. pp. 173–176 (2006)