An Interactive Visual Analytic Tool for Semantic Classification of 3D Urban LiDAR Point Cloud

Beena Kumari and Jaya Sreevalsan-Nair

Abstract

We propose a novel unsupervised machine learning approach for effective semantic labeling by combining two different multi-class classifications, structural and contextual classification, of points in airborne LiDAR point cloud of urban environment. Structural classification labels a point in the cloud as a point-, line-, or surface-type feature. An additional outcome of this classification is the geometry-preserving downsampling of the point cloud. The contextual classification, on the other hand, labels the points in four classes, namely, buildings, vegetation, natural ground, and asphalt ground, by using data derived from the raw input, which includes the structural classification. Preserving these two classifications in the labeling of the points gives a geometry-aware contextual semantic labeling. We propose: (a) an augmented semantic classification which preserves both structural and contextual classification, (b) an interactive hierarchical clustering method for contextual classification, and (c) an interactive visual analytic framework to aid both the structural and contextual classifications.

Keywords: Semantic Classification, 3D LiDAR Point Cloud Visualization, Visual Analytics, Feature Extractions,

I. Introduction

Airborne LiDAR is a widely used technology for collection of topographic data. Extraction of spatial features in LiDAR datasets of urban environment is known as semantic labeling or classification. Though there exist several supervised or unsupervised classification techniques for semantic labeling, the domain expert or human-in-the-loop who analyzes the data or generates the training set are indispensable to ensure that the points are correctly classified. At the same time, the functional need of urban planners, spatial data users, and engineers is to explore and analyze the various features of these datasets, especially new ones, for an initial evaluation, where a domain expert may not be necessary.

Our aim is to devise a solution for two use case scenarios. Firstly, it is helpful for a LiDAR data user to do an initial coarse classification of new datasets, which can potentially even aid in improving data acquisition. However he/she is unable to do classification without the help of a domain expert. Secondly, in the current approaches, a user or even a domain expert does not have the flexibility to explore a new dataset and determine a classification technique based on the relevant features in that particular dataset. Our proposed solution is an interactive visual analytic framework for urban LiDAR point cloud using unsupervised techniques, where the classification is interactively determined by the user.

We show an implementation of such a tool that can be used to extract structural as well as contextual features and combine them together to give an augmented classification, in comparison to the existing methods. The novelties of our semantic classification, which we refer to as *augmented semantic classification*, are: (a) it includes an interactive hierarchical (top-down) clustering, hence is unsupervised, and (b) it is geometry as well as contextual aware. In our framework, user can interactively visualize the point cloud, detect the structural features, visualize the various feature descriptors derived from the datasets, and make informed decisions on parameters based on supporting visualizations for structural as well as contextual classification.

II. Related Work

There is a large body of work in the area of semantic classification of LiDAR data. Our visual analytic framework for the augmented semantic classification, and the interactive divisive clustering for contextual classification are novel ideas, and we have not found any prior relevant work on those lines. We discuss some of the work which is relevant to various modules of our proposed framework. Blomley et al. [2] have used multi-scale approach using shape distribution features for point classification, as opposed to covariance features, proposed by Keller et al. [4]. We have used the latter in our work. Niemeyer et al. [8] have used conditional random field with random forest classifier to include context to exploit spatial locality. In similar lines, we refer to our object-based classification as a contextual one.

The contextual classification of real world datasets is not an exact science, as it is subjective to the topography, data acquisition, the heterogeneity of materials (man-made, natural, vegetation, asphalt, etc.), LiDAR parameters, and other metadata of the dataset. Hence, point-wise local descriptors are required at each point for classification. Tombari et al. [12] have categorized local descriptors into signature and histogram, and this categorization has been used implicitly in our proposed method. Using both types of local descriptors have been proven to give better segmentation of the point cloud [7], [9]. In our proposed work, we use signature descriptors for detection of structural features for structural classification; and in the case of spatial features for contextual classification, we use a combination of descriptors, which includes both signature and histogram descriptors.

III. Our Proposed Method

Our proposed method comprises of two modules: (a) augmented semantic classification (Section III-C), and (b) visual analytic framework which enables visualization and data analytics in an interleaved manner (Section III-D). Our augmented semantic classification preserves both structural and contextual classifications, discussed in Sections III-A, and III-B, respectively. Our proposed system is summarized in Figure 1, where the GUI is built upon our visual analytic framework and the data is preprocessed before the classifications.
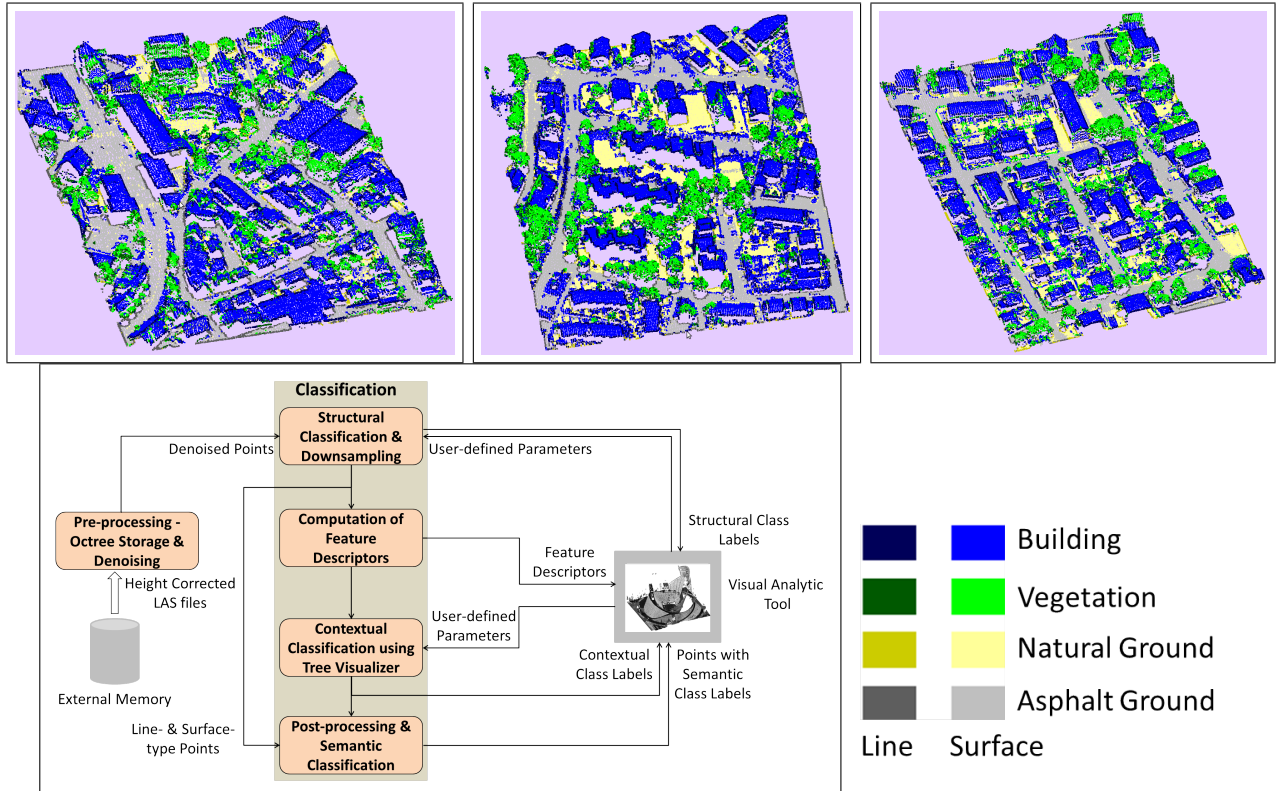
Fig. 1. Our augmented semantic classification of Areas 1, 2, and 3 (left-to-right) of the Vaihingen dataset [3], on the top; and our proposed visual analytic framework to achieve the augmented semantic classification.

A. Structural Classification

Raw LiDAR data contains noise due to instrumental errors and atmospheric conditions in which data has been collected. Therefore, denoising of data has been done using conditional outliers removal filter [5]. LiDAR dataset has ellipsoidal height as elevation value, not a normalized one which is a very important feature descriptor. Therefore, normalized height has been computed using LAStools[1].

We have used the algorithm proposed by Keller et al. [4] for structural classification. It classifies the points in the cloud as disc, curve, critical curve, and critical disc points. The disc and curve points refer to points on surfaces and lines (such as creases, borders, and ridges), respectively, and hence, are surface- and line-type features, respectively. The critical curve and critical disc points refer to point-type feature. This algorithm has been referred to as structural feature-, geometry-, or topology-based, owing to the eigen-analysis of the covariance matrix of the local neighborhood, which gives the local reference frame (RF) [12]. The local RF computed using eigenvalue decomposition of the covariance matrix has been proven to be one of most robust local RF [12]. The likelihood of the points falling into any of the three classes (i.e., point, line, surface) is determined by saliency values, C_p, C_l, C_s, respectively, which are computed using a multi-scale approach. Since the saliency values are computed from local RF, they are signature descriptors.

Another beneficial outcome of this algorithm is downsampling, which is possible by retaining features, more persistent across scales. The downsampling ensures that the topology of the point cloud is preserved and hence, there is no loss of salient information. Downsampled LiDAR dataset is easier to store, analyze, and visualize. Since multi-scale implementation is not usually preferred due to its computationally expensive nature [2], we alleviate the problem by using the parallel implementation, proposed by Kumari et al. [5].

B. Contextual Classification

We perform contextual classification on the downsampled point cloud, using interactive hierarchical divisive clustering to obtain four object classes. By interactive hierarchical clustering, we imply that the user can customize the parameters used for clustering done at each level or subtree in the hierarchy. Our algorithm, in particular, is divisive clustering, as, at each node in the k-ary tree of the clustering hierarchy, the user decides the subset of feature descriptors, to be used as parameters for splitting the set of points described by the node. The user makes an informed decision of the clustering parameters after observing clustering tendency of the feature descriptors using supporting visualizations. Any efficient partitioning clustering algorithm, such as Expectation Maximization (EM), k-means, graph cuts, etc., can be used to split the clusters, once the user has confirmed that there are clear clusters for the chosen set of points in the chosen parameter space. We have used the following combination of feature descriptors for our classification:
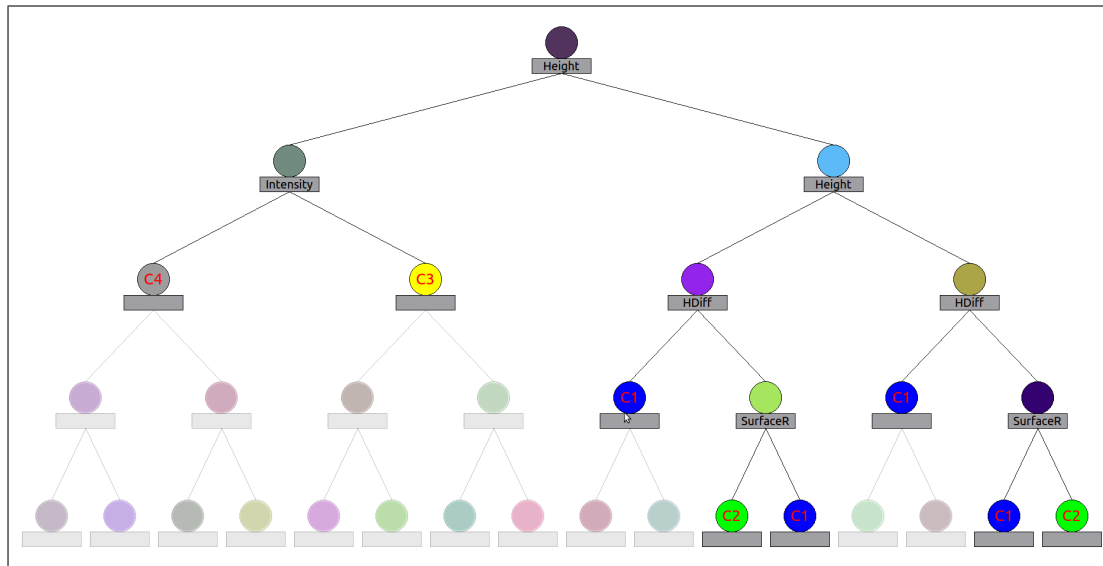

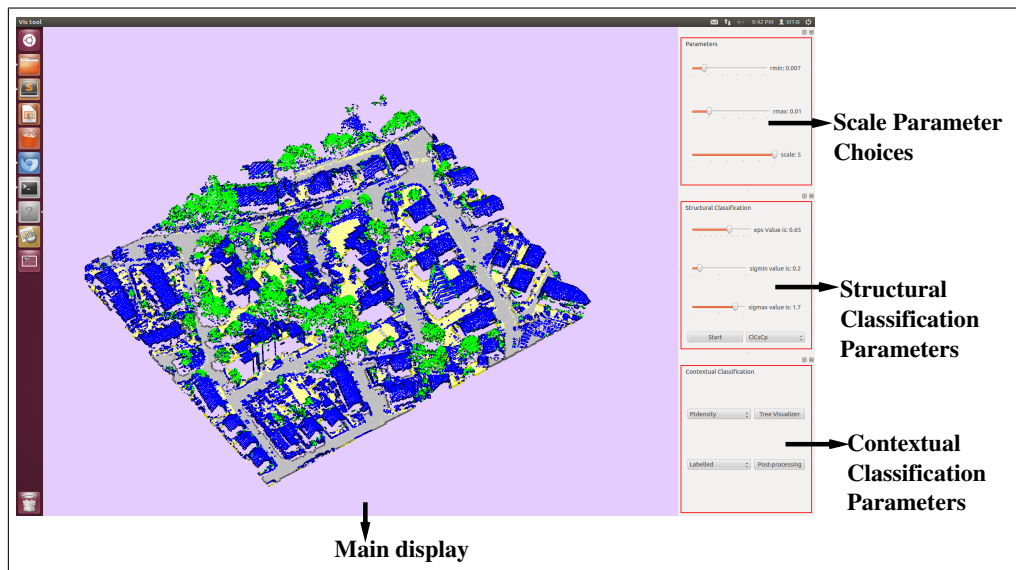[1]http://rapidlasso.com/lastools/lasheight/

Fig. 2. (Top) The main window of our Graphical User Interface (GUI) for our visual analytic framework, showing main display on the left, and widget controls on the right, for scale parameters, structural classification, and contextual classification (top-to-bottom). (Bottom) GUI for our tree visualizer for hierarchical clustering. In the active leaf nodes of our tree visualizer, (which are not dimmed/transparent), the blue ones indicate building, yellow natural ground, green vegetation, and gray asphalt ground. The points belonging to different clusters of the same class (e.g. blue clusters) get labeled the same (e.g. building).

1) Height (h): is normalized height, i.e., height of objects above the ground, and is good for separating ground and non-ground points.
2) Intensity (I): is sensitive to the incidence angle of the laser beam [8] and we have used I to distinguish the asphalt ground from natural ground.
3) Height variance (σ_h^2): is the variance of the normalized height computed from $r-$nearest neighbors, in cylindrical neighborhood, and is known to be high for vegetation and gabled roofs.
4) Height range (δ_h): is the difference between maximum and minimum normalized height of $r-$nearest neighbors, in cylindrical neighborhood, and is known to be high for vegetation and boundary of buildings.
5) Eigen-based LiDAR features - are computed using the eigen values, $\lambda_1, \lambda_2, \lambda_3$, of covariance matrix computed using $r-$nearest neighbors, in spherical neighborhood (different from other descriptors). Anisotropy ($\mathcal{A}$), a feature, is known to be high for planar region and low for non-planar region, and conversely, sphericity ($\mathcal{S}$), another feature, is known to be high for non-planar region and low for planar regions; and hence can be used to distinguish vegetation from planar roofs. Given $\lambda_1 \geq \lambda_2 \geq \lambda_3$, $\mathcal{A} = \frac{\lambda_1 - \lambda_3}{\lambda_1}$, and $\mathcal{S} = \frac{\lambda_3}{\lambda_1}$.
6) Surface fit (SF): tells how best the local surface fits at any point p of point cloud using $r-$nearest neighbors, in spherical neighborhood and is known to be high for planar regions and low for non-planar. It can be used for distinguishing planar roof from vegetation.

Items 3 and 4 can be considered to be histogram based while items 5 and 6 are signature descriptors; together they are local feature descriptors, whereas those in item 1 and 2 are point-based. In our clustering hierarchy, the downsampled point

cloud is the root node, and at each subsequent level, we split it into k clusters using EM Gaussian mixture model (GMM). While the number of clusters for splitting at each node in the tree may be a variable, for our application, $k = 2$ throughout the tree was optimal for efficient computations for three-fold reasons: (a) clustering results can be updated interactively, (b) large-scale datasets can be handled, and (c) finding two clear clusters in the parameter space at each split is easier than larger number of clusters. Furthermore, we have performed an intra-class region-growing [11] for points belonging to the building and vegetation classes, exclusively. This gives us segments, also referred to as features, which essentially preserves the spatial context.

C. Augmented Semantic Classification

The motivation for preserving the geometric features is that identifying lines in the point cloud is beneficial for retrieving building footprints or road bank boundaries. The lines in the classification outcome are currently computed using: (a) bounding polygon of clustered building and road class points, respectively [1], or (b) using LiDAR elevation data [6]. However, the methods which use bounding polygons solely depend on accuracy of classification, and those which superimpose imagery and point cloud, face various problems [1]. The problems in latter method include: (a) due to non-availability of accurately co-registered imagery and LiDAR data, there exist unresolved differences in LiDAR data and imagery, as both are collected separately, and (b) due to shadow and occlusion, some information is not clearly available in imagery. Thus, it is challenging to combine the feature descriptors from data collected in two different modes. Hence, our idea is to preserve and use the line-type points, computed from a single mode, i.e., the point cloud itself. Additionally, the geometric features, computed as described in Section III-A, are relevant local feature descriptors computed from a robust local RF, which makes the case for preserving these specific features. Our proposed augmented semantic classification aims to preserve geometric as well as contextual information that each point in the point cloud embodies. Thus, each point in the downsampled cloud has a label, which is an ordered pair of geometric and contextual class labels.

D. Visual Analytic Framework

A visual analytic framework enables user interactivity for doing data mining, visualization, as well as selection of subsets of datasets, for further exploration and analysis. Our proposed framework has the following design requirements:

D1. The user must be provided with information on the various feature descriptors to make decisions for the interactive hierarchical clustering.
D2. The user must be able to change parameters for both classifications, and be able to visualize the results interactively, so that an optimum combination of parameters can be arrived at for any new dataset.
D3. The user must be able to explore meaningful subsets of the point cloud, i.e., points in the same class or cluster given by the structural or contextual classification.

In our framework, heatmaps of various descriptors are provided catering to D1; the tree visualizer for hierarchical divisive clustering (Figure 2) catering to D2; class visualization to focus on points in selected classes catering to D3. The visual analytic tool, as shown in Figure 2, comprises of the main display of point clouds, tree visualizer, a group of widgets for multi-scale setting, another group for parameter tuning for structural classification, a drop-down menu for heatmaps of feature descriptors, a drop-down menu for visualizing subsets of point cloud belonging to specific structural classes, and tab for opening files.

IV. EXPERIMENTS & RESULTS

We have tested our algorithm on the following ISPRS benchmark airborne LiDAR datasets: Areas 1, 2, and 3, in Vaihingen test site [3]. Table I gives the augmented semantic classification of the test datasets. The %age of points in each of the semantic classes are comparable across the datasets, even though the number of objects, say in the building class for the benchmark data given in [10], are not similar. This anomaly could be attributed to the non-uniform point density in the different classes caused by the downsampling. The percentage reduction in Area 1, Area 2 and Area 3 are 37.67, 31.91 and and 31.97 respectively. Hausdorff distance between original and downsampled point clouds for Areas 1, 2, and 3, have been computed to be 9.4170m, 8.3800m, and 10.5077m, respectively, which is negligible given the spread of the sites, as given in Table I. The geometry-awareness of our augmented segmented classification can be clearly seen in Figure 1, where we can see the points in (line, building) and (line, asphalt ground) classes showcase delineation of the building boundaries and road banks.

C-class → / S-class ↓	Building	Vegetation	Natural Ground	Asphalt Ground
	(in % points)			
Area 1 (200m × 250m, 2.12 Pts/m^2)				
Surface	30.74	10.68	6.81	20.14
Line	12.04	8.44	3.53	7.61
Area 2 (225m × 256m, 2.76 Pts/m^2)				
Surface	26.58	10.45	13.42	20.29
Line	12.92	6.53	4.25	5.43
Area 3 (193m × 277m, 2.54 Pts/m^2)				
Surface	27.43	9.55	13.68	21.34
Line	11.61	5.60	5.32	5.44

TABLE I
AUGMENTED SEMANTIC CLASSIFICATION SHOWING STRUCTURAL(S-CLASS) AND CONTEXTUAL(C-CLASS) CLASSES ALONG ROWS, FOR LiDAR BENCHMARK DATASETS.

Our tool has been tested by domain experts in geographic information system (GIS), visualization, and machine learning. The GIS expert mentioned that the attractiveness of tool is the compelling use case of identifying line features and overlaying

it on the object-based classification; however commented on the requirement to scale up the application to larger areas. He mentioned that the accuracy of the footprints generated from our tool in comparison to the corresponding orthoimages can be visually rated at 80-85%. The visualization expert commented on the effective use of heatmaps for deciding on parameters for contextual classification, and the possibility of extending the visual analytics capability to study covariance of parameters taken pairwise. The machine learning expert commented on the limitations of using EM algorithm owing to the underlying data model not being directly representable as a Gaussian mixture model, and encouraged us to explore the effects of changing the clustering algorithm.

V. Conclusions

We have proposed a technique for augmented semantic classification of LiDAR point cloud data, by combining structural and contextual classification. Our method uses a multi-scale approach to preserve geometric features, and the structural classification additionally downsamples the point cloud, in a geometry-preserving manner. A new unsupervised machine learning approach, namely, an interactive hierarchical divisive clustering algorithm followed by region growing has been proposed for contextual classification. A visual analytic framework has been implemented for augmented semantic classification as a proof-of-concept. Our results show that contextual classification gives visually accurate results, and the geometry-awareness of our augmented semantic classification has benefits in visualizing and identifying boundaries and structure. Domain experts who have reviewed our tool have appreciated the usefulness of interactivity in determining the hierarchical clustering, and the interactive updates of the visualizations, owing to the geometry-preserving downsampling and parallel implementation.

VI. Acknowledgements

References

[1] Mohammad Awrangjeb and Clive S Fraser. Automatic segmentation of raw lidar data for extraction of building roofs. *Remote Sensing*, 6(5):3716–3751, 2014.

[2] R Blomley, M Weinmann, J Leitloff, and B Jutzi. Shape distribution features for point cloud analysis-a geometric histogram approach on multiple scales. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1:9–16, 2014.

[3] Michael Cramer. The dgpf-test on digital airborne camera evaluation–overview and test design. *Photogrammetrie-Fernerkundung-Geoinformation*, 2010(2):73–82, 2010.

[4] Patric Keller, Oliver Kreylos, Marek Vanco, Martin Hering-Bertram, Eric S Cowgill, Louise H Kellogg, Bernd Hamann, and Hans Hagen. Extracting and visualizing structural features in environmental point cloud LiDaR data sets. In *Topological Methods in Data Analysis and Visualization*, pages 179–192. Springer, 2011.

[5] Beena Kumari, Avijit Ashe, and Jaya Sreevalsan-Nair. Remote Interactive Visualization of Parallel Implementation of Structural Feature Extraction of Three-dimensional Lidar Point Cloud. In *Big Data Analytics*, pages 129–132. Springer, 2014.

[6] Xiaoye Liu. Airborne LiDAR for DEM generation: some critical issues. *Progress in Physical Geography*, 32(1):31–49, 2008.

[7] J Niemeyer, F Rottensteiner, and U Soergel. Conditional random fields for lidar point cloud classification in complex urban areas. *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences*, 1(3):263–268, 2012.

[8] Joachim Niemeyer, Franz Rottensteiner, and Uwe Soergel. Contextual classification of lidar data and building object detection in urban areas. *ISPRS journal of photogrammetry and remote sensing*, 87:152–165, 2014.

[9] AM Ramiya, Rama Rao Nidamanuri, and R Krishnan. Semantic Labelling of Urban Point Cloud Data. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1:907–911, 2014.

[10] Franz Rottensteiner, Gunho Sohn, Jaewook Jung, M Gerke, Caroline Baillard, Sebastien Benitez, and Uwe Breitkopf. The isprs benchmark on urban object classification and 3d building reconstruction. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences I-3*, pages 293–298, 2012.

[11] Radu Bogdan Rusu and Steve Cousins. 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)*, China, 2011.

[12] Federico Tombari, Samuele Salti, and Luigi Di Stefano. Unique signatures of histograms for local surface description. In *Computer Vision–ECCV 2010*, pages 356–369. Springer, 2010.