Segmentation and Localization of Individual Trees from MMS Point Cloud Data Acquired in Urban Areas

MARTIN WEINMANN[1,2], CLÉMENT MALLET[1] & MATHIEU BRÉDIF[1]

Abstract: In this paper, we address tree segmentation and localization in the scope of the IQmulus Processing Contest IQPC'15. Based on the part of pre-classified 3D point cloud data which corresponds to trees, we present a novel framework which involves a downsampling of the original data, a projection of the downsampled data onto a horizontally oriented plane, a mean-shift-based segmentation of the projected points, a transfer of the segmentation results to the original data, a refinement of the segmentation results via segment-based shape analysis, and a localization of respective tree trunks. The results derived for a benchmark dataset reveal that all individual trees are correctly detected and localized with both acceptable accuracy and reasonable computational effort.

Zusammenfassung: In diesem Beitrag wird die Segmentierung und Lokalisierung von Bäumen im Rahmen des IQmulus Processing Contest IQPC'15 behandelt. Basierend auf jenem Anteil von vorab klassifizierten 3D-Punktwolken, welcher Bäume repräsentiert, wird eine Methodik vorgestellt, die eine teilweise Betrachtung der Originaldaten, eine 2D-Projektion entsprechender Punkte auf eine horizontal ausgerichtete Ebene, eine Segmentierung der projizierten Punkte basierend auf dem Mean-Shift-Verfahren, einen Transfer der Segmentierungsergebnisse zu den Originaldaten, eine Verbesserung der Segmentierungsergebnisse basierend auf einer Analyse der Segmentform und eine Lokalisierung entsprechender Baumstämme umfasst. Die für einen Referenz-Datensatz erreichten Ergebnisse zeigen auf, dass alle darin enthaltenen Bäume mit ausreichender Genauigkeit und angemessenem Rechenaufwand erkannt und lokalisiert werden können.

1 Introduction

The automated analysis of 3D point cloud data is of great interest in photogrammetry and remote sensing. Amongst a variety of research topics, the detection of individual trees from 3D point cloud data has been paid particular attention in recent years. In the context of forest applications, for instance, airborne laser scanning data has been used at tree-level to delineate individual crowns, to estimate wood volumes and biomass, or to identify the respective tree species. In contrast, mobile mapping systems allow a ground-based acquisition of significantly denser 3D point cloud data which are meanwhile used to derive tree cadasters in urban environments, e.g. for applications in city modeling and city planning.

In the IQmulus Processing Contest IQPC'15, point cloud analysis has recently been addressed with a special track on tree separation and classification (GORTE et al. 2015). Thereby, the focus is on 3D point cloud data acquired with mobile mapping systems and, hence, mainly data

[1] Université Paris-Est, IGN, SRIG, MATIS, 73 avenue de Paris, 94160 Saint-Mandé, France, E-Mail: [martin.weinmann, clement.mallet, mathieu.bredif]@ign.fr

[2] Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing, Englerstraße 7, 76131 Karlsruhe, Germany, E-Mail: martin.weinmann@kit.edu

acquired along roads and in urban environments. While the first step consists in a classification into 3D points corresponding to trees and other 3D points, the second step consists in separating those 3D points corresponding to the "tree" class into clusters referring to the respective individual trees (Fig. 1).

Since the emphasis of IQPC'15 is on tree separation, we focus on a straightforward and efficient approach for individual tree segmentation and localization in the scope of this paper, so that scalability towards the processing of even larger datasets is given. In contrast to approaches relying on a voxelization of 3D space to reduce the computational burden, we present an approach which considers the data on point-level (which is highly relevant for an accurate delineation of individual trees and also for tree parameter extraction) but increases computational efficiency by conducting a typically rather time-consuming mean-shift-based segmentation only for a specific subspace as well as a downsampled version of the original data. The obtained segmentation results are then transferred to the original data. A subsequent refinement of the segmentation results is carried out via a segment-based shape analysis and yields plausible tree segments. Finally, the tree location is estimated by detecting a respective tree trunk. We demonstrate the performance of our approach for a recent benchmark dataset acquired in an urban environment, and we discuss the achieved results with regard to competing approaches.

The paper is structured as follows. After briefly discussing related work in Section 2, we present our proposed methodology for tree segmentation and localization in Section 3. Subsequently, we provide experimental results in Section 4 and discuss these results in detail in Section 5. Finally, in Section 6, we provide concluding remarks and suggestions for future work.

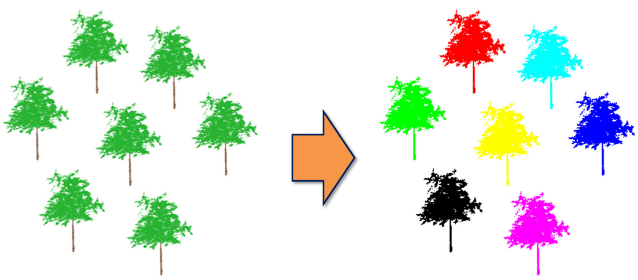

Fig. 1: Schematic illustration for the segmentation of single trees from all 3D points classified as tree points. All those 3D points not belonging to the "tree" class are assumed to have been removed by the previous step of tree classification.

2 Related Work

Point cloud segmentation aims to provide a meaningful partitioning of a set of 3D points into smaller, connected subsets which correspond to objects of interest or to parts of these (MELZER 2007; VOSSELMAN 2013). In the context of our work, we deal with 3D point cloud data acquired with a mobile mapping system, and the derived segments should represent individual trees.

Many approaches for detecting single trees from 3D point cloud data rely on a voxelization of 3D space. In YAO & FAN (2013), it is for instance proposed to derive a 2D accumulation map on a horizontally oriented plane and – based on respective features – to separate natural objects such as trees from man-made objects. Those 3D points corresponding to natural objects are transferred to a voxel space and then a normalized cut segmentation based on the voxel structure is carried out (REITBERGER et al. 2009). A different strategy for detecting and segmenting individual trees consists in first classifying laser scanning point clouds into "tree" and "non-tree" classes (SIRMACEK & LINDENBERGH 2015), and subsequently focusing on tree individualization. The latter may be achieved by a voxelization of 3D space, deriving connected components and separating the components further if they contain multiple clusters (GORTE et al. 2015). Alternatively, tree individualization may be achieved by a downsampling and retiling of the original 3D point cloud data via voxelization, where a subsequent 2D gridding allows to find local maxima in point density and thus potential tree locations (LINDENBERGH et al. 2015). Based on these tree locations, individual trees are finally segmented via octree-based region growing and thresholding techniques.

Instead of involving a voxelization of 3D space, it has been proposed to perform tree individualization by considering the original data on point-level and calculating geometric descriptors for each 3D point, projecting these descriptors onto a horizontally oriented 2D accumulation map and considering a spatial filtering to obtain individual tree clusters (MONNIER et al. 2012). A further approach for tree individualization on point-level focuses on deriving connected components for those 3D points categorized into a "tree" class, and the connected components are further split via an upward and downward growing algorithm if there are multiple seeds at a height between 0.5m and 1m (GORTE et al. 2015; OUDE ELBERINK & KEMBOI 2014). Furthermore, a direct consideration of the original data on point-level is possible by applying a standard clustering technique such as *k*-means clustering or hierarchical clustering (GUPTA et al. 2010), or the mean shift algorithm presented in (FUKUNAGA & HOSTETLER 1975). The latter has for instance been applied on 3D point cloud data in (FERRAZ et al. 2012; SCHMITT et al. 2013; YAO et al. 2013; SHAHZAD et al. 2015). However, such an approach can be computationally demanding, particularly for a large number of considered 3D points. To improve computational efficiency, it seems desirable to apply the mean shift algorithm on a 2D projection of a 3D point cloud as e.g. described in (SCHMITT et al. 2015) for 3D point cloud data acquired via tomographic SAR processing. However, the point density will be significantly higher for a ground-based acquisition of 3D point cloud data as e.g. given when using mobile mapping systems. Hence, handling the amount of available data still remains challenging due to the computational burden arising from data processing and data storage.

3 Methodology

Our proposed framework for tree segmentation and localization is illustrated in Fig. 2. It consists of 6 successive steps represented by (1) a downsampling of the original data, (2) a projection of the downsampled data onto a horizontally oriented plane, (3) a mean-shift-based segmentation of the projected points, (4) a transfer of the segmentation results to the original data, (5) a refinement of the segmentation results via segment-based shape analysis, and (6) a localization of respective tree trunks. The input of the framework is given by 3D points which – in a previous step – have been classified to correspond to a tree (Fig. 3), and the output of the framework is given by individual tree segments and the location of respective tree trunks. More details on the single components are provided in the following subsections.

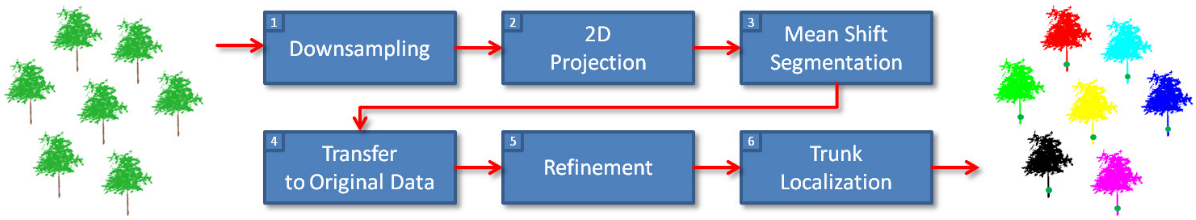

Fig. 2: The proposed framework for segmenting and localizing individual trees from 3D point cloud data acquired with a mobile mapping system: the input data are downsampled and projected onto a horizontally oriented plane; a mean-shift-based segmentation is subsequently carried out in 2D space and results are transferred to the original data; a shape-based refinement yields plausible tree segments and respective trunks are finally localized for these plausible segments.

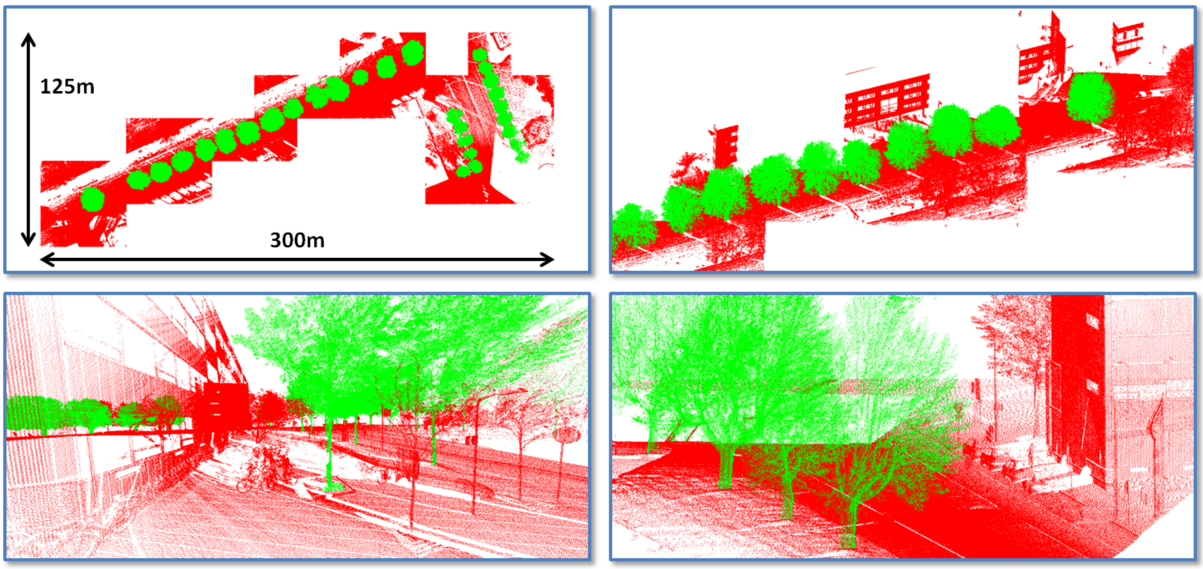

Fig. 3: Pre-classified 3D point cloud data (top row: nadir view and side view; bottom row: more detailed views): 3D points categorized into the “tree” class are indicated in green and serve as input for our framework, whereas all other 3D points are indicated in red and not considered by our framework.

3.1 Downsampling of the Original 3D Point Cloud

Taking into account that MMS point cloud data provides a very dense representation of object surfaces near the acquisition system, it might be desirable to conduct time-consuming tasks only for a downsampled version of the original data to reduce the computational effort with respect to processing time and memory consumption. In this regard, it also becomes obvious that the point density of 3D points corresponding to trees may significantly be decreased while still being able to detect individual trees in the respective 3D point cloud data. Accordingly, we conduct a downsampling by only keeping every k-th point of the original 3D point cloud. In the results, we will report results for k = 10, a parameter selected heuristically. Alternatively, a pruning of this parameter could be conducted based on the local point density (CARAFFA et al. 2015).

3.2 Projection Onto Horizontally Oriented Plane

In contrast to forested areas, urban areas typically provide a larger spacing and less overlap between individual trees due to human intervention in nature and different planning processes. Accordingly, we propose to neglect the occurrence of dominant, co-dominant or dominated trees in urban environments and assume that individual trees may still sufficiently be delineated when only considering a 2D projection of the downsampled 3D point cloud data corresponding to trees onto a horizontally oriented plane.

3.3 Segmentation via Mean Shift Algorithm

In order to derive a meaningful partitioning of the downsampled 3D point cloud data, we consider the respective 2D projections on the horizontally oriented plane and apply the mean shift algorithm (FUKUNAGA & HOSTETLER 1975; CHENG 1995; COMANICIU & MEER 2002; MELZER 2007) for segmentation. Generally, the mean shift algorithm represents an iterative, statistical technique for locating the maxima / modes of a probability density function by only considering discrete data sampled from that probability density function. There is no need to explicitly estimate the underlying probability density function itself, and the mean shift algorithm does neither rely on the assumption of a specific geometric model nor require an initial definition of the number of modes.

In our case, we treat the 2D projections of the downsampled 3D point cloud data as discrete data points sampled from an empirical 2D probability density function. For each data point, the mean shift algorithm iteratively (1) calculates the weighted mean of data points within a window defined by a kernel K (typically an isotropic kernel such as a Gaussian kernel or an Epanechnikov kernel (COMANICIU & MEER 2002)), (2) defines the mean shift vector as the difference between the data point and the weighted mean of data points within the considered window, and (3) moves the data point along the mean shift vector. In this context, the step-size defined by the length of the mean shift vector will be large in areas of low point density, whereas it will be low in areas of high point density. In this way, the mean shift algorithm iteratively performs an adaptive gradient ascent until convergence (up to numerical accuracy). The stationary points correspond to regions of high point density and represent the modes of the underlying distribution of data points. Finally, all data points leading to the same mode are considered as cluster or segment. For our application, the single clusters / segments are expected to represent the individual trees in the considered scene.

The only parameters to be specified are represented by the bandwidth parameters indicating the size of the kernel in each direction. Due to the consideration of data points on the horizontally oriented plane, we may assume the same bandwidth h in all directions, which is intuitively justified since we want to detect individual trees. Furthermore, we may involve prior knowledge about the shape and size of the trees in the considered scene and hence heuristically select the bandwidth parameter. Based on different tests, we heuristically select a bandwidth parameter of h = 3.8m for the involved dataset.

3.4 Transfer of Segmentation Results to Original Data

To get from the segmentation results for the downsampled version of the original data back to the original 3D point cloud data, we consider an intuitive, simple and straightforward approach. This approach assigns each 3D point of the original 3D point cloud data the segment label of the closest 3D point in the downsampled version of the original data. Thereby, we conduct a nearest neighbor search based on Euclidean distances. As a result, we have transferred the segmentation results from the downsampled data to the original 3D point cloud data.

3.5 Refinement of Segmentation Results

The previous steps lead to a coarse segmentation of the considered 3D point cloud data, yet they do not guarantee that the derived segments are plausible. For our application, we want to have meaningful segments corresponding to individual trees. It becomes obvious that – for a segment corresponding to an individual tree – the top parts of the segment are characterized by foliage, whereas the lower parts are characterized by the tree trunk. We hence focus on checking each derived segment for a trunk-like structure at its lower parts. For this purpose, we consider the lowest 3D point of the respective segment as basis and cut the part up to the diameter of breast height (which may vary between 1.3m and 1.5m depending on the tree type and depending on the country) into horizontally oriented slices of a certain width. More specifically, we select 6 slices covering the lowest 1.5m of each segment, i.e. each slice has a width of 0.25m. Since circle fitting tends to be impractical because of only partly visible trunks and the fact that possibly occurring trunk radii would have to be pre-defined for a model-based approach, we focus on an approximate data-driven solution. We first calculate the mean value m_i of all 2D coordinates per slice i and then define the mean value M across these means as potential location of the tree trunk. Subsequently, we take into account the standard deviation σ of the distance between slice-wise means m_i and the potential location M of the tree trunk. If the standard deviation σ becomes too large (e.g. $\sigma > 0.2$m) or if the derived clusters comprise a relatively low number N_p of points (e.g. $N_p < 100$ points), we assume that the respective segment does not correspond to a tree. In this case, the segment is either merged with the next-closest plausible segment in case of directly adjacent segments or, otherwise, removed from the set of 3D points which have been classified to correspond to a tree.

3.6 Localization of Tree Trunks

In the last step, we consider each plausible segment for an individual tree and define the location of the corresponding tree trunk to coincide with the respectively estimated mean position M across the 6 lowest slices defined for that segment.

4 Experimental Results

The involved benchmark dataset has been acquired in the vicinity of the campus of TU Delft in the Netherlands with the Fugro DRIVE-MAP system (GORTE et al. 2015). For our experiments, we consider the provided subset comprising 26 tiles, for which results of a previous binary classification are available (Fig. 3). The 26 tiles contain a total number of about 10.13M points of which about 1.78M points (17.6%) are assigned to the "tree" class, and we intend to subdivide the set of all these 3D points into segments corresponding to individual trees (i.e. we address the Challenge 1 of the IQPC'15 special track on tree separation and classification).

For our experiments, we conduct a downsampling by only keeping every 10th point of the original 3D point cloud. Then, we perform the 2D projection and apply the mean shift algorithm with a bandwidth parameter of h = 3.8m. The results obtained for the benchmark dataset are shown in Fig. 4. With our approach, we have found a total number of 29 trees in the scene and a visual inspection reveals that all the derived segments correspond to individual trees. Thus, the segmentation of individual trees is reached with an accuracy of 100%. A more detailed analysis of the derived results reveals that, due to the localization of trees based on considering the lowest slices of plausible segments, all estimated tree locations coincide with a respective tree trunk. Note that, for the considered benchmark dataset, these estimates are biased towards the position of the sensor, since only a part of the tree trunks is acquired with the mobile mapping system. However, we consider the estimate to be sufficiently accurate for our application.

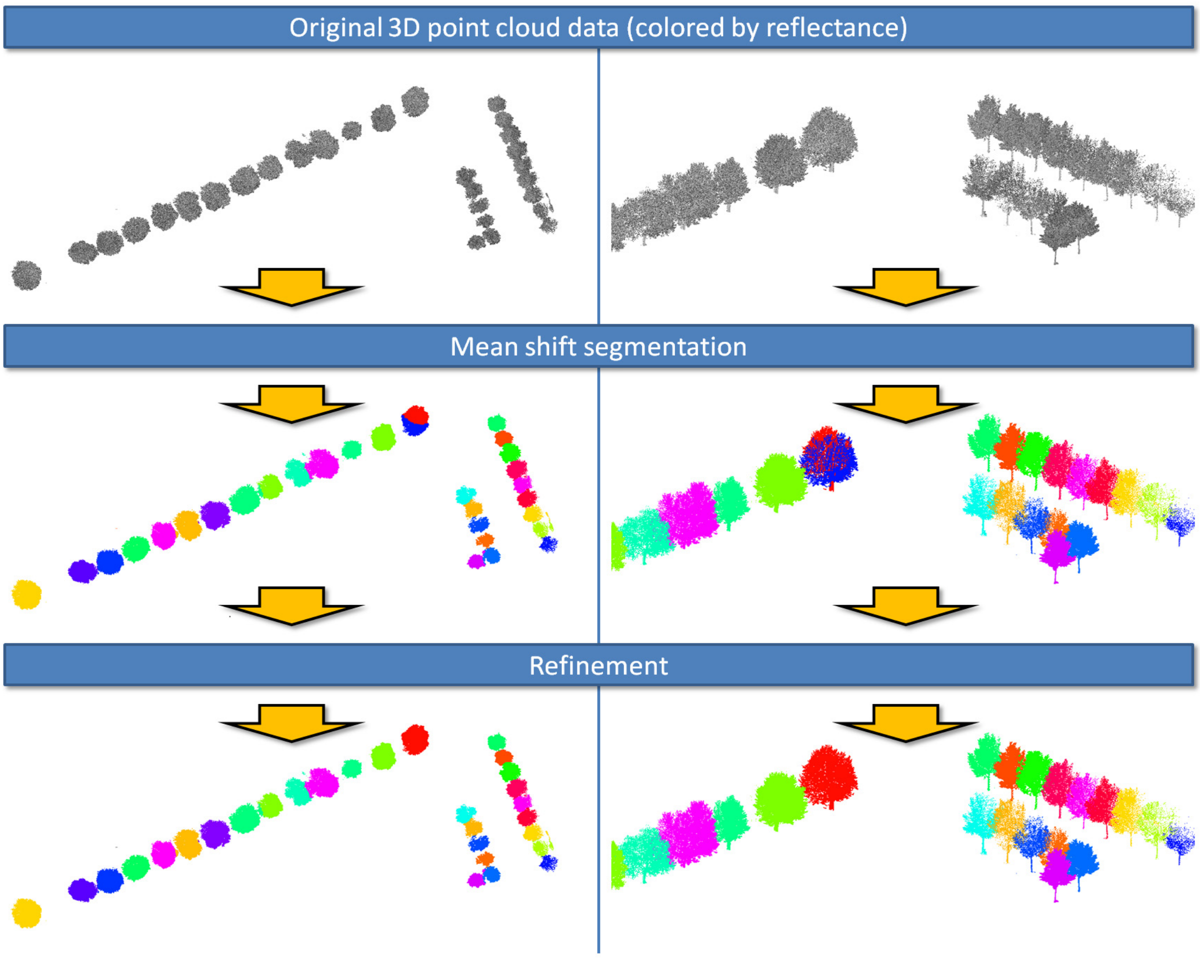

Fig. 4: The involved dataset in nadir view (left column) and in a side view (right column): original 3D point cloud data (top row), segmentation result after applying the mean shift algorithm on a 2D projection of a downsampled version of the original data and the transfer back to the original data (center row), and segmentation result after the shape-based refinement (bottom row).

The required computational effort for the whole methodology comprising downsampling, 2D projection, mean shift segmentation, transfer of segmentation results to original data, refinement and trunk localization is less than 1 minute on a notebook of medium performance (Intel Core i5-2410M, 2.3GHz, 4GB RAM, Matlab implementation).

5 Discussion

As already pointed out in (MELZER 2007) in the context of point cloud segmentation, one of the most important advantages of applying the mean shift algorithm consists in its capability to directly work on the original data. Since this may however be rather time-consuming for large datasets, we apply the mean shift algorithm only on a downsampled version of the original data and subsequently transfer the segmentation results back to the original data. Thereby, we further involve a 2D projection, since the computational effort of a mean shift segmentation in 2D is significantly lower than for a mean shift segmentation in 3D (FERRAZ et al. 2012; SCHMITT et al. 2013). Thus, our strategy reduces the computational burden considerably without significantly affecting the quality of the derived segmentation results.

Besides representing an efficient approach for individual tree segmentation and localization, our approach is rather simple and easy-to-use. It directly works on the given data without relying on a voxelization as e.g. presented in (GORTE et al. 2015; LINDENBERGH et al. 2015), where the voxel size as well as the voxel orientation might strongly influence the respective segmentation results. Considering the derived segmentation results, we may state that our results are sufficiently accurate for the benchmark dataset and that only minor segmentation errors at segment borders are visible if adjacent trees are relatively close to each other. However, the latter also becomes visible in the results presented in (GORTE et al. 2015).

A disadvantage of the proposed approach consists in the fact that it relies on one critical parameter represented by the bandwidth parameter h which has a strong influence on the number of clusters. However, since it has a physical meaning, we have selected a suitable value by involving heuristic knowledge about the expected size and shape of trees in the scene which is in accordance to (SCHMITT et al. 2015). Furthermore, it is obvious that the results of the presented approach strongly depend on the quality of the previous separation between 3D points corresponding to trees and 3D points corresponding to other objects, but this also holds for the competing approaches presented in (GORTE et al. 2015; LINDENBERGH et al. 2015).

6 Conclusions & Future Work

In this paper, we presented a framework for tree segmentation and localization representing an important topic in the scope of the IQmulus Processing Contest IQPC'15. The input to our framework consists in 3D points which – in a previous step – have been classified to correspond to a tree. Our framework is composed of (1) a downsampling of the original data, (2) a projection of the downsampled data onto a horizontally oriented plane, (3) a mean-shift-based segmentation of the projected points, (4) a transfer of the segmentation results to the original data, (5) a refinement of the segmentation results via segment-based shape analysis, and (6) a localization of respective tree trunks. The results derived for a benchmark dataset reveal that all individual

trees are correctly detected and localized with acceptable accuracy and reasonable computational effort. Only minor segmentation errors may be observed at the boundaries of some segments, but these errors are in the same order as in related investigations (GORTE et al. 2015).
For future work, we intend to address the fundamental requirement for our framework, i.e. the appropriate separation of 3D points belonging to trees from 3D points belonging to other objects. Particularly the use of geometric features e.g. as the ones presented in (WEINMANN et al. 2015a,b) seems to be promising for this task, since respective geometric features allow an appropriate separation of 3D points corresponding to foliage, trunk and ground (STUCKER 2015).

7 Acknowledgements

This work was supported by the European Commission's Seventh Framework Programme under grant agreement FP7-ICT-2011-318787 (*IQmulus: A High-Volume Fusion and Analysis Platform for Geospatial Point Clouds, Coverages and Volumetric Data Sets*).

8 References

CARAFFA, L., BRÉDIF, M. & VALLET, B., 2015: 3D octree based watertight mesh generation from ubiquitous data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40** (3/W3), 613-617.

CHENG, Y., 1995: Mean shift, mode seeking, and clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence **17** (8), 790-799.

COMANICIU, D. & MEER, P., 2002: Mean shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (5), 603-619.

FERRAZ, A., BRETAR, F., JACQUEMOUD, S., GONÇALVES, G., PEREIRA, L., TOMÉ, M. & SOARES, P., 2012: 3-D mapping of a multi-layered Mediterranean forest using ALS data. Remote Sensing of Environment **121**, 210-223.

FUKUNAGA, K. & HOSTETLER, L., 1975: The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Transactions on Information Theory **21** (1), 32-40.

GORTE, B., OUDE ELBERINK, S., SIRMACEK, B. & WANG, J., 2015: IQPC 2015 Track: Tree separation and classification in mobile mapping lidar data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40** (3/W3), 607-612.

GUPTA, S., WEINACKER, H. & KOCH, B., 2010: Comparative analysis of clustering-based approaches for 3-D single tree detection using airborne fullwave lidar data. Remote Sensing **2** (4), 968-989.

LINDENBERGH, R.C., BERTHOLD, D., SIRMACEK, B., HERRERO-HUERTA, M., WANG, J. & EBERSBACH, D., 2015: Automated large scale parameter extraction of road-side trees sampled by a laser mobile mapping system. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40** (3/W3), 589-594.

MELZER, T., 2007: Non-parametric segmentation of ALS point clouds using mean shift. Journal of Applied Geodesy **1** (3), 159-170.

MONNIER, F., VALLET, B. & SOHEILIAN, B., 2012: Trees detection from laser point clouds acquired in dense urban areas by a mobile mapping system. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **1** (3), 245-250.

OUDE ELBERINK, S. & KEMBOI, B., 2014: User-assisted object detection by segment based similarity measures in mobile laser scanner data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40** (3), 239-246.

REITBERGER, J., SCHNÖRR, CL., KRZYSTEK, P. & STILLA, U., 2009: 3D segmentation of single trees exploiting full waveform lidar data. ISPRS Journal of Photogrammetry and Remote Sensing **64** (6), 561-574.

SCHMITT, M., BRÜCK, A., SCHÖNBERGER, J. & STILLA, U., 2013: Potential of airborne single-pass millimeterwave InSAR data for individual tree recognition. Publications of the German Society for Photogrammetry, Remote Sensing and Geoinformation, Volume 22, Potsdam, Seyfert, E. (Ed.), 33. Wissenschaftlich-Technische Jahrestagung der DGPF, 27 February - 1 March 2013, Freiburg, 427-436.

SCHMITT, M., SHAHZAD, M. & ZHU, X. X., 2015: Reconstruction of individual trees from multi-aspect TomoSAR data. Remote Sensing of Environment **165**, 175-185.

SHAHZAD, M., SCHMITT, M. & ZHU, X. X., 2015: Segmentation and crown parameter extraction of individual trees in an airborne TomoSAR point cloud. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XL** (3/W2), 205-209.

SIRMACEK, B. & LINDENBERGH, R., 2015: Automatic classification of trees from laser scanning point clouds. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **2** (3/W5), 137-144.

STUCKER, C., 2015: Klassifizierung von Vegetation in einer Laserscan-Punktwolke. Publications of the German Society for Photogrammetry, Remote Sensing and Geoinformation, Volume 24, Potsdam, Kersten, T.P. (Ed.), 35. Wissenschaftlich-Technische Jahrestagung der DGPF, 16-18 March 2015, Cologne, 370-375.

VOSSELMAN, G., 2013: Point cloud segmentation for urban scene classification. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **40** (7/W2), 257-262.

WEINMANN, M., JUTZI, B., HINZ, S. & MALLET, C., 2015a: Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. ISPRS Journal of Photogrammetry and Remote Sensing **105**, 286-304.

WEINMANN, M., URBAN, S., HINZ, S., JUTZI, B. & MALLET, C., 2015b: Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas. Computers & Graphics **49**, 47-57.

YAO, W. & FAN, H., 2013: Automated detection of 3D individual trees along urban road corridors by mobile laser scanning systems. Proceedings of the International Symposium on Mobile Mapping Technology (MMT) 2013, 1-6.

YAO, W., KRZYSTEK, P. & HEURICH, M., 2013: Enhanced detection of 3D individual trees in forested areas using airborne full-waveform lidar data by combining normalized cuts with spatial density clustering. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **2** (5/W2), 349-354.