POINT CLOUD SEGMENTATION FOR URBAN SCENE CLASSIFICATION

George Vosselman

Faculty ITC, University of Twente, Enschede, the Netherlands – george.vosselman@utwente.nl

KEY WORDS: Segmentation, Classification, Point cloud, Urban, Airborne, Filtering

ABSTRACT:

High density point clouds of urban scenes are used to identify object classes like buildings, vegetation, vehicles, ground, and water. Point cloud segmentation can support classification and further feature extraction provided that the segments are logical groups of points belonging to the same object class. A single segmentation method will typically not provide a satisfactory segmentation for a variety of classes. This paper explores the combination of various segmentation and post-processing methods to arrive at useful point cloud segmentations. A feature based on the normal vector and flatness of a point neighbourhood is used to group cluttered points in trees as well as points on surfaces in areas where the extraction of planes was not successful. Combined with segment merging and majority filtering large segments can be obtained allowing the derivation of accurate segment feature values. Results are presented and discussed for a 70 million point dataset over a part of Rotterdam.

1. INTRODUCTION

The classification of point clouds is an important step in the extraction of information. Whereas point cloud classification initially served to select points on the ground in the context of DTM production, the higher point densities obtained nowadays allow the extraction of various object types. This requires classification with further classes like buildings, vegetation, vehicles, and water. Algorithms for point cloud classification are typically either point-based or segment-based. Point-based approaches, unlike the name may suggest, not just use make use of point features like reflectance strength or echo count, but also of features derived from a point's neighbourhood like height variation or RMS values of local plane fitting. The key characteristic is that these features are separately calculated for every point and that every point is separately classified. Context can be taken into account by point-based methods through the use of probabilistic relaxation (Smeeckaert et al., 2013) or graphical models like conditional random fields (Niemeyer et al., 2012). Without the use of context large height variations observed around building edges can e.g. easily be interpreted as a characteristic of vegetation and lead to misclassification (Vosselman et al., 2004).

A segment-based approach classifies segments of the point cloud based on segment features. These features may be determined by e.g. averaging over feature values of the points of a segment, but may also be specific to the segment, like e.g. the size or shape descriptors. The latter features are an extension to the set of point-based features and may improve the discrimination between classes. The integration of point feature values over segments may also lead to more accurate feature values and thereby improve classification results. Key to the success of a segment-based classification is, of course, the segmentation. In the case of under-segmentation, points of different classes will be part of the same segment. As all points of a segment will obtain the same class label, any under-segmentation will lead to classification errors. Lim and Suter (2009) therefore on purposely over-segment a point cloud before classifying segments with a conditional random field. Over-segmentation, however, reduces the quality of the segment features. Integration over smaller amounts of point feature values will lead to less noise reduction. Furthermore, segment shape descriptors may become less useful.

The advantages and disadvantages of a segment-based point cloud classification are very similar to those of segmented-based image classification, commonly called object-based image analysis (OBIA) (Hay and Castilla, 2006). Although efforts are made to optimise parameter settings for multi-resolution image segmentation (Dragut et al., 2010), there is little discussion on the segmentation methods themselves.

A single point cloud segmentation method will typically not provide a satisfactory segmentation. This paper explores the combination of various segmentation and post-processing methods to arrive at useful point cloud segmentations that contain possibly many points per segment to derive accurate attribute values and at the same time minimise under-segmentation. A brief review on segmentation method is provided in section 2. Many segmentation methods are designed to extract surfaces, e.g. for the extraction of terrain pieces or roof faces. These methods are not suitable to capture objects like trees, poles, and (depending on the point density) vehicles. The review therefore emphasises methods that do not focus on the extraction of (planar) surfaces. Section 3 describes the drawbacks of segmentation into planes and the use of connected components. The combination of multiple segmentation methods is considered necessary. In Section 4 a feature is described to allow grouping of points in vegetation as well as grouping of points on a surface. Post-processing methods to increase the size of segments and thereby make their features more representative are discussed in Section 5. Throughout this paper examples are shown from laser scanning survey conducted over Rotterdam with a point density of 30 points/m^2.

2. RELATED LITERATURE

Many algorithms have been developed for the extraction of surfaces from point clouds. Efficient RANSAC (Schnabel et al., 2007) and 3D Hough transform combined with surface growing (Vosselman and Klein, 2010) are often used in work to extract

roof faces and other surfaces from airborne laser scanning data. Less studied are methods to segment point clouds of objects that are not necessarily described in terms of surfaces. Melzer (2007) presented a first study to apply mean shift (Comaniciu and Meer, 2002) to the segmentation of urban point clouds. Points on buildings, vegetation and terrain were already grouped by using mode seeking with only the X-. Y-and Z-coordinates. Finer segmentations were obtained when also making use of amplitude and pulse width of the echoes. Ferraz et al. (2010) used mean shift to separate surface vegetation, understory and overstory in forested areas. Yao et al. (2009) combined mean shift with normalised cuts to extract vehicles and flyovers. Rutzinger et al. (2009) used segment growing to cluster and classify vegetation points in an urban environment. Only the homogeneity in echo widths was used as a criterion for clustering neighbouring points. This feature typically distinguishes vegetation from smooth surfaces. Some over-segmentation in vegetation was observed because of variation in the echo widths within the vegetation.

More work on segmenting point clouds into non-planar segments has been performed with mobile laser scanning data. A typical workflow is to determine the points on ground surface, remove those points from the dataset and then determine the connected components in the remaining point set (Douillard et al., 2010). Pu et al. (2011) and Velizhev et al. (2012) in addition incorporated scene knowledge to select components for further classification. Pu et al. eliminated large vertical components (walls) when extracting street furniture whereas Velizhev et al. selected on component size and distance to the ground when selecting cars and street lights. Golovinskiy and Funkhouser (2009) made initial estimates of background points (street level) and foreground points (street furniture, cars) and then used a min-cut based segmentation to improve the initial estimates.

Aijazi et al. (2013) segmented a point cloud generated by mobile laser scanning in two steps. After removing points on the ground the remaining connected components are segmented based on colour and reflectance strength. Another two-step approach has been presented by Xu et al. (2012). After an initial segmentation and classification of planar point sets, connected components of points with a doubtful classification were re-segmented using mean shift to generate new segments for a further classification.

3. SEGMENTATION IN MULTIPLE STAGES

In this section we discuss the advantages and limitations of segmentation into planes (3.1) and the use of connected components (3.2) for segmenting airborne laser scanning data.

3.1 Segmentation into surfaces

Figure 1 shows a segmentation of a point cloud of an urban area obtained by growing planar segments from seeds detected by a 3D Hough transform. Roof faces are well captured in segments. Most walls are also extracted as planar segments although the point density on walls is clearly lower. To obtain this result the neighbourhood used for growing was defined by the k nearest neighbours without a restriction on the distance between a point and its neighbours.

Vegetation is split into many small planar segments and points not belonging to any planar segment (white points in Figure 1). The planar segments do not represent surface parts of the trees, but are sets of points on different branches that are nearly coplanar. The higher the point density and the larger the point-

to-plane tolerance, the higher the likelihood will be that a set of arbitrary points in vegetation is considered co-planar. The sizes of those segments are typically very small. Hence, the segment size can be a useful feature to distinguish vegetation from roof faces (Xu et al., 2012). Other features, however, become inaccurate because of the low number of points in a segment. E.g., the percentage of echoes not being the last echo, which is typically high in vegetation, but low on roofs and other surfaces, will be unreliable for classification of small vegetation segments.

Another artefact in the segmentation is the fragmentation of the terrain surface. As the terrain is not exactly planar, the street surface breaks up into nearly co-planar larger surfaces. Features describing the point distribution within these segments may be affected by the segments' seemingly arbitrary shapes.
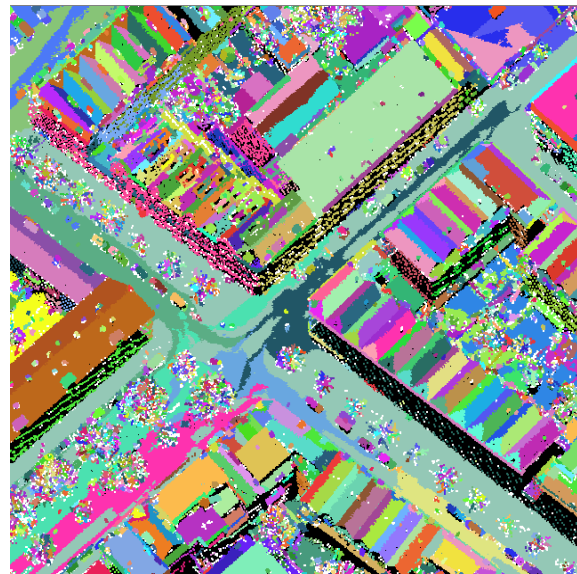

Figure 1. Segmentation of a point cloud into planar point sets.

3.2 Connected components of unsegmented points

As described in section 2, processing of mobile laser scanning data often groups points on street furniture and cars by a connected component analysis after removing the points on the street level. A similar approach can be applied to the airborne laser scanning data. After keeping apart all larger segments (more than 100 points in the example) connected components can be determined in the remaining point set.

Figure 2 shows the result for the example of Figure 1. Points on trees and cars typically form clear segments. Some nearby trees and nearby cars are merged. As long as the merged objects belong to the same class, this will not have an effect on the classification accuracy unless shape descriptors are used as part of the segment features.

The initial segmentation will also have contained correctly detected small surfaces, like roofs of dormer windows. These will not have been kept apart, but will have been re-segmented by the connected component analysis. Typically, all points on a dormer window are then again grouped to a segment.

While the connected components discussed so far are useful as they group points of the same object class, Figure 2 also shows some larger segments that combine pieces of vegetation, smaller patches of ground points and smaller pieces of walls. Such segments will inevitably lead to classification errors. Although the percentage of points in such mixed segments is relatively small in the example (about 1%), a classification of a larger segment of e.g. vegetation points to wall points may lead

to locally very disturbing errors. To avoid such errors the connected components need to be split up further.

This then leads to the strategy to segment point clouds in multiple stages. First, the larger planar segments are extracted as the points in those segments typically correspond to the same class. Smooth surfaces that break up into multiple planar surfaces may be merged in a post-processing step. After removing the larger planar segments from the point set, the remaining points are again segmented, but now with different criteria such that points in vegetation may group together.
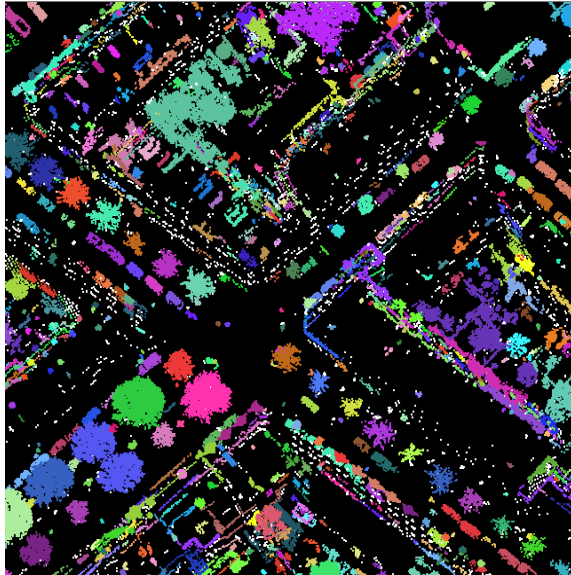

Figure 2. Connected components of points belonging to planar segments of less than 100 points.

4. SEGMENTING NON-PLANAR COMPONENTS

4.1 Criteria for homogeneity

Larger connected components typically contain combinations of vegetation, terrain and walls. To separate those classes use can be made of the local point cloud planarity as derived from the coordinate co-variance matrix in a point's neighbourhood in combination with the normal vector direction. The planarity distinguishes the vegetation from the terrain and walls, whereas the normal vector direction is used to separate terrain from wall points. When $\lambda_1 > \lambda_2 > \lambda_3$ are the eigenvalues of the co-variance matrix, the planarity can be expressed as $(\lambda_2 - \lambda_3) / \lambda_2$. In the methods discussed below, the planarity was combined with the normal vectors by a multiplication. Points in vegetation have more or less random vector directions. By multiplying the vectors with the typically low planarity value, the resulting scaled vectors cluster in the feature space around the null vector. For walls and terrain patches the planarity value will be close to one. Consequently, a multiplication of the normal vectors with planarity values results in a feature space in which points of vegetation, walls, and terrain will be clustered at different locations.

In literature, planarity is often defined as $(\lambda_2 - \lambda_3) / \lambda_1$ (Chehata et al., 2010). This, however, leads to low planarity values in case of points on elongated pieces of wall for which $\lambda_1 \gg \lambda_2$. Normalisation by λ_2 is therefore preferred.

Planarity should also be preferred over anisotropy $((\lambda_1 - \lambda_3) / \lambda_1)$. Although anisotropy is generally high in vegetation and low for ground and wall surfaces, point clouds of larger trees often show points grouped on branches. For such linear point distributions the value of λ_3 is low compared to λ_1, i.e. the anisotropy is high. As λ_2 and λ_3 are often similar (branches are round), the planarity value is still low. Hence, a low planarity value is obtained for cluttered as well as linear point clouds and therefore suitable for the recognition of trees with larger branches as a single segment.

4.2 Experiments with Mean shift

Melzer (2007) used mean shift to segment point clouds based on only the coordinates or together with amplitude and pulse width extracted from full waveforms. Just using the coordinates did not lead to separation of the different classes of points for the used dataset. When adding the scaled normal vector elements as features a better separation was obtained. Vegetation in gardens was, however, often split in multiple segments without a clear spatial separation. I.e., some points of segment A were found in the middle of points of segment B and vice versa. This comes as the result of the mix of coordinate differences and feature value differences in the multivariate kernel function. When feature values are very similar, points may be grouped into the same segment despite a slightly larger Euclidian distance between the points. Balancing the bandwidths of coordinates and other feature values cannot completely avoid this characteristic of the mean shift segmentation. As a consequence, the distribution of points within a single segment produced by mean shift segmentation can be rather inhomogeneous. This makes it more difficult to characterise the segments with features based on point distributions, like e.g. coordinate variances or point density. Hence, the obtained segments seem less suitable for a segment based classification.

4.3 Segment growing

To ensure a good spatial coherence of the points belonging to one segment, a segment growing algorithm was used. Instead of testing neighbouring points on the distance to a plane as commonly done for surface growing, the test for accepting neighbouring points as extensions of a segment is now based on
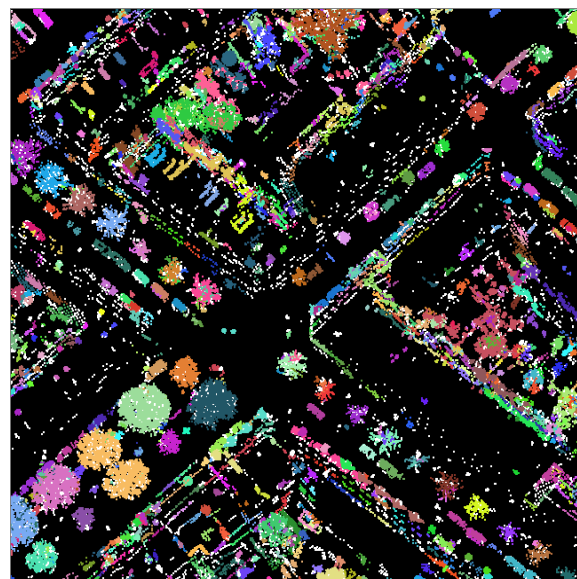

Figure 3. Segment growing based on normal vectors scaled by planarity.

similarity of feature values. Rutzinger et al. (2009) used such segment growing based on echo widths of full waveforms. The echo width will typically separate vegetation from non-vegetation and in that sense fulfils a similar function as the local point cloud planarity. To also distinguish between surfaces with different orientations, again the normal vectors scaled by the planarity were used. Seeds for the growing were defined as sets of neighbouring points with similar feature values. These seeds were then extended with adjacent points if their feature values are close to the average feature values of the segment.

The result of such a segmentation is shown in Figure 3. In general most points in trees were grouped together. Trees, however, also show some white points. These are points with deviating feature values that did not have sufficient nearby similar points to start a new seed. A simple post-processing step to include these points in the segmentation is described in section 5.2.

The large connected components in Figure 2 break up into smaller segments after the segment growing. Only some smaller segments contain points from multiple classes. Hence, the classification errors that will be caused by under-segmentation are strongly reduced.

For the estimation of the planarity a neighbourhood of 50 points was used. Clearly smaller neighbourhoods, e.g. 20 points, have an increased change that points in vegetation show a close to planar distribution.

Using a neighbourhood size of 50 points implies that all points in connected components of 50 points or smaller will have exactly the same feature values. As a consequence, such components will not be further segmented by the segment growing. Smaller objects like dormer windows and chimneys will therefore be detected as a single segment, just like they were determined by the connected component analysis.

In principle one could also apply the segment growing to the original point cloud without first extracting and removing the larger planar segments. Experiments, however, showed that the quality of such segmentations in inferior to the segmentation in two stages because the surface growing is dedicated to the extraction of planes and more accurately extracts the point sets that are truly planar.

5. POST-PROCESSING

To improve the usefulness of the obtained segments for a segment based classification two post-processing steps are applied. In the first step (nearly) co-planar segments are merged. In the second step isolated points without a segment number are assigned to a segment.

5.1 Merging planar segments

As shown in Figure 1 smooth but non-planar surfaces will be split up into multiple planar patches. Shape and point distribution features of such segments may then lead to an incorrect classification. To avoid this two neighbouring segments will be merged if they are nearly co-planar at their common border.

This merging of segments is also important for handling large datasets. As large point clouds cannot be dealt with in memory at once, point clouds are typically tiled and processed tile by tile. In urban areas points on an object will often be distributed over multiple tiles, e.g., when a tile boundary intersects a roof. In particular when a tile boundary is close to an object's edge, it will split off a small part of the object. The resulting segment will then be difficult to classify due to inaccurate feature values

and a lack of context from surrounding segments. To reduce this problem Xu et al. (2012) used tiles with an overlap. This provided some more context, but still insufficient to deal with the classification of water surfaces. Figure 4 shows classification results with rectangular patches of ground (grey) surrounded water (blue). These patches correspond to the tiles used for processing the data. Xu et al. (2012) used the point density within a segment to discriminate between terrain and water surfaces. For some of the tiles in the middle of a strip the point density on the water was very high and led to a misclassification. The classification was further complicated by the changing water levels in this harbour area. Surfaces in different strips therefore had different heights. As a result some surface patches above the "ground" (lower patches) were classified as building roof (shown in red).
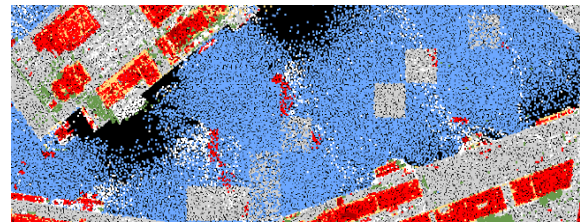

Figure 4. Classification errors in water surfaces.

Merging nearly co-planar segments within tiles and across tile boundaries resolves the above described problems. Figure 5 shows the segmentation results of nine tiles before and after segment merging. The segment classification can now benefit from more accurate feature values and more context information.
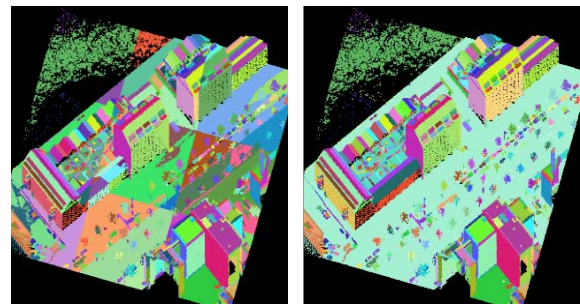

Figure 5. Segment merging within tiles and across tile boundaries. Left: segmentation results per tile. Right: merged segments.

Figure 6 shows the segmentation results on a 70 million point dataset over a part of Rotterdam. The area of 2.4 km^2 was split into 960 tiles of 50x50 m. After segment merging the whole road network including the adjacent ground surfaces becomes one single segment of 30 million points. As the road smoothly connects to the bridge surface, the latter is also included in the large terrain segment.

The water surfaces corresponding to the different flight lines are now also well recognisable. Surfaces with more than 20 cm in between were extracted separately. The large water segments now combine data of tiles with high point densities with data of tiles with low point densities. The presence of areas with low point densities within a segment is a strong indicator for a water surface. The large water segments can therefore be clearly distinguished from the large ground segments. This will solve the classification problems shown in Figure 4.
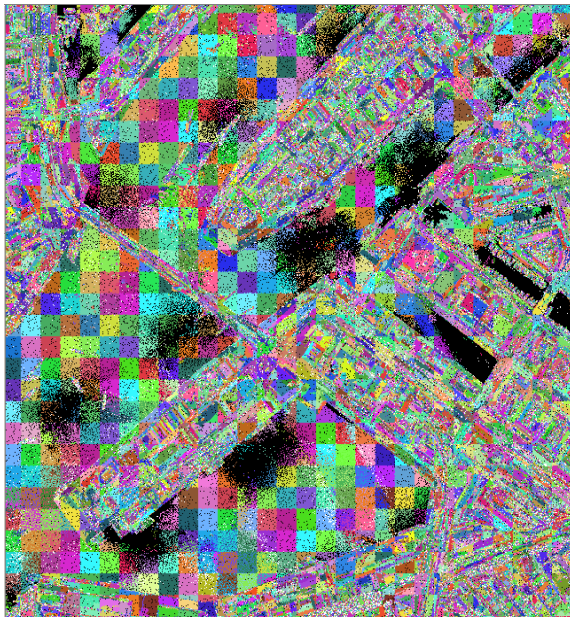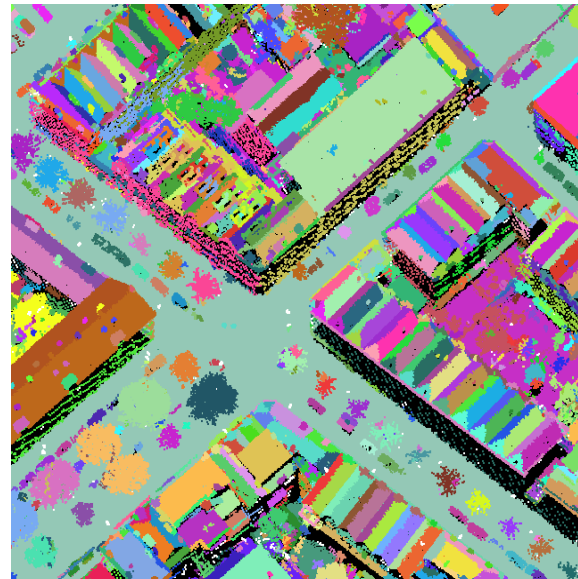
Figure 7. Final segmentation results.

The combined results in Figure 7 show that nearly all points belong to larger segments. Most trees are recognised as a single segment. Some trees, however, were segmented into two or three parts. The feature values of those segments are nevertheless still quite different that those of segments of other classes. A few nearby trees and cars are merged.

6. DISCUSSION

By combining different segmentation and post-processing methods a segmentation result can be obtained which groups points on vegetation, ground, walls, roofs, and water surfaces. Only a few segments contain points of multiple classes. As most segments contain many points feature values can be computed that will be representative for the class of a segment. This will likely increase the quality of a segment-based classification.

The features used for segmentation should be calculated with an appropriate neighbourhood size. As Brodu and Lague (2012) showed the neighbourhood size in vegetation will decide whether vegetation is observed as clutter (many branches and leaves in one large neighbourhood), planar (co-planar branches or single leaves) or linear (small neighbourhood with only single branches or twigs). The choice of suitable features for segmentation (and classification) therefore also depends on the point density.

As discussed in section 2 efforts to extract and classify street furniture from mobile laser scanning data typically use connected components of points above the street level as the units for classification. The large variety of shapes of traffic lights, street lights, and other objects may make it difficult to classify street furniture based on features of the point sets. When high point density is available, it would be recommendable to further segment the connected components and try to recognise parts like straight and curved cylinders and planar pieces. Such a decomposition may lead to a richer description and better classification of street furniture.

The described segmentation approach clearly is making use of specific knowledge on the object classes to be recognised; most roof faces are planar, terrain is a mostly smooth surface, and vegetation appears as a clutter. Although the segmentation algorithms do not classify the points, the assignment of a point

Figure 6. Segment merging in a larger data set. Top: segmentation results per tile. Bottom: merged segments.

5.2 Majority filtering

The segment growing results showed a large number of isolated points that did not merge into any segment. To assign these points to segments a majority filter is used. The most frequent segment number within a fixed radius neighbourhood of an unsegmented point is assigned to this point. This proofs to be an effective way to obtain large vegetation segments.

All segmentation and post-processing steps lead to the segmentation result in Figure 7. After the extraction of planar segments (3.1) nearly co-planar segments were merged (5.1). All points in segments smaller than 100 points and all points without segment numbers were re-segmented by segment growing based on the point cloud planarity and normal vector direction (4.3). Finally, non-segmented points were labelled by majority filtering if there are neighbouring points within some radius.

to a segment clearly has a large impact on the later classification of that point. A further integration of segmentation and classification was shown in Xu et al. (2012) and Van Den Eeckhaut et al. (2012) where an initial segment-based classification formed the context for a further segmentation and classification phase. Whereas these studies only used two segmentation phases (the latter only applied to a small part of the data), an even tighter integration of segmentation and classification may lead to further improvements in the execution of both tasks.

REFERENCES

Aijazi, A.K., Checchin, P., Trassoudaine, L., 2013. Segmentation Based Classification of 3D Urban Point Clouds: A Super-Voxel Based Approach with Evaluation. *Remote Sensing* 5, 1624-1650.

Brodu, N., Lague, D., 2012. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: Applications in geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing* 68, 121–134.

Chehata, N., Guo, L., Mallet, C., 2010. Airborne lidar feature selection for urban classification using random forests. In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, part 3/W8, Paris, France, 1-2 September, pp. 207-212.

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.

Douillard, B., Underwood, J., Vlaskine, V., Quadros, A., Singh, S., 2010 A Pipeline for the Segmentation and Classification of 3D Point Clouds. *Proceedings of the International Symposium on Experimental Robotics* (ISER 2010).

Dragut, L., Tiede, D., Levick, S. R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24 (6), 859-871.

Ferraz, A., Bretar, F., Jacquemoud, S., Goncalves, G., Pereira, L., 2010. 3D segmentation of forest structure using a Mean-shift based algorithm. Proceedings IEEE 17th International Conference on Image Processing, 26-29 September, Hong Kong, pp. 1413-1416.

Golovinskiy, A., Funkhouser, T., 2009. Min-Cut Based Segmentation of Point Clouds. *Workshop at the IEEE 12th International Conference on Computer Vision*, 27 September – 4 October, pp. 39-46.

Hay, G.J., Castilla, G., 2006. Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). In: *The International Archives on Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 36, Part 4/C42, 3p.

Lim, E., Suter, D., 2009. 3D terrestrial lidar classifications with super-voxels and multi-scale conditional random fields. *Computer-Aided Design* 41(10), 701–710.

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

Niemeyer, J., Rottensteiner, F., Soergel, U., 2012. Conditional random fields for lidar point cloud classification in complex urban areas. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Science*, Melbourne, Australia, Vol. I-3, pp. 263-268.

Pu, S., Rutzinger, M., Vosselman, G. and Oude Elberink, S.J., 2011. Recognizing basic structures from mobile laser scanning data for road inventory studies. *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (6 Supplement), S28-S39.

Rutzinger, M., Höfle, B, Hollaus, M., Pfeifer, N., 2008. Object-Based Point Cloud Analysis of Full-Waveform Airborne Laser Scanning Data for Urban Vegetation Classification. *Sensors* 2008, 8, 4505-4528.

Schnabel, R., Wahl, R., Klein, R., 2007. Efficient RANSAC for point-cloud shape detection. *Computer Graphics Forum* 26 (2), 214-226.

Smeeckaert, J., Mallet, C., David, N., Chehata, N., Ferraz, A., Large-scale water classification of coastal areas using airborne topographic lidar data. *Proceedings of the 33rd IEEE International Geoscience and Remote Sensing Symposium* (IGARSS 2013).

Van Den Eeckhaut, M., Kerle, N., Poesen, J., Hervás, J., 2012. Object-oriented identification of forested landslides with derivatives of single pulse LiDAR data. *Geomorphology* 173–174, 30–42.

Velizhev, A., Shapovalov, R., Schindler, K., 2012. Implicit shape models for object detection in 3D point clouds. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Science*, Melbourne, Australia, Vol. I-3, pp. 179-184.

Vosselman, G., Gorte, B.G.H., Sithole, G., 2004a. Change detection for updating medium scale maps using laser altimetry. In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 34, Part B3, Istanbul, Turkey, 12-23July, pp. 207-212.

Vosselman, G., Klein, R., 2010. Visualisation and structuring of point clouds. In: Vosselman, G., Maas, H.-G. (Eds.), *Airborne and Terrestrial Laser Scanning*, Whittles Publishing, pp. 45-81.

Xu, S., Oude Elberink, S.J. and Vosselman, G., 2012. Entities and Features for Classification of Airborne Laser Scanning Data in Urban Area. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Science*, Melbourne, Australia, Vol. I-4, pp. 257-262.

Yao, W., Hinz, S., Stilla, W., 2009. Object extraction based on 3d-segmentation of LiDAR data by combining mean shift with normalized cuts: two examples from urban areas. In: *Proceedings of 2009 Joint Urban Remote Sensing Event (URBAN2009 - URS2009)*, Shanghai, China.