CLASSIFICATION OF MOBILE LASER SCANNING POINT CLOUDS FROM HEIGHT FEATURES

Mingxue Zheng[a,b], Mathias Lemmens[b], Peter van Oosterom[b]

[a] State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing,
Wuhan University, Wuhan, China
[b] Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, the Netherlands
mingxuezheng@tudelft.nl, M.J.P.M.Lemmens@tudelft.nl, P.J.M.vanOosterom@tudelft.nl

Commission WG II/3

KEY WORDS: Point clouds, Mobile Laser Scanning, Feature extraction, Classification, Urban area, Vertical objects

ABSTRACT:

The demand for 3D maps of cities and road networks is steadily growing and mobile laser scanning (MLS) systems are often the preferred geo-data acquisition method for capturing such scenes. Because MLS systems are mounted on cars or vans they can acquire billions of points of road scenes within a few hours of survey. Manual processing of point clouds is labour intensive and thus time consuming and expensive. Hence, the need for rapid and automated methods for 3D mapping of dense point clouds is growing exponentially. The last five years the research on automated 3D mapping of MLS data has tremendously intensified. In this paper, we present our work on automated classification of MLS point clouds. In the present stage of the research we exploited three features – two height components and one reflectance value, and achieved an overall accuracy of 73%, which is really encouraging for further refining our approach.

1. INTRODUCTION

Point clouds acquired by MLS are attribute poor. In addition to the 3D coordinates in a local, national or regional reference system, usually only the reflectance value of each point is available in a point cloud data set. As a result, many classification approaches rely on enriching the attribute set with RGB values from imagery, which may not always be available, and on examining the local geometric structure of a set of neighbouring points. The latter approach is based on the observation that many objects differ in shape. There are two main basic descriptors available to express the shape of a surface: normal vectors and eigenvalues. Both are assigned to individual points by examining the configuration of the point and its adjacent points. The normal vectors are computed per point by reconstructing the surface sampled by the point itself and the points in its vicinity. If the normal vectors of neighbouring points point in the same direction, the local neighbourhood forms likely a plane. If they diverge in a systematic manner they likely form a sphere or a cylinder. When no systematics in directions is present, the points may be reflected on a fuzzy surface, such as foliage. Also, eigenvalues of the 3x3 covariance matrix of the three coordinates of neighbouring points indicate shape. If one eigenvalue is large and the other two close to zero, the neighbourhood forms a line. A plane is indicated by two eigenvalues which have approximately the same value and one eigenvalue close to zero. Spherical and fuzzy surfaces will have three large eigenvalues. It is common practice to derive measures from the eigenvalues which indicate the type of local structure. Examples of such measures are: linearity, planarity, sphericity, anisotropy, eigenentropy and local surface variation. Methods based on the exploitation of the local geometric structure have been studied by multiple researchers. Recent overviews are provided by Guan et al. (2016), Grilli et al. (2017) and Lemmens (2017).

Indeed, because point clouds have only few attributes it is inevitable to explore a local neighbourhood in the class assignment process of individual points. We propose a method which differs from the ones discussed above. In the case of 3D mapping of outdoor scenes, the heights above a reference surface, e.g. ground surface, are the most important asset of a point cloud and this information should by fully exploited (Lemmens, 2017). In a previous tentative research which aimed at exploring the height component we used the height above ground level (Zheng et al., 2017). It appeared that the height value of a point does not discriminate enough among the different classes. Points reflected on façades, cars, motorcycles, pedestrians and traffic signs all may have the same height. In this paper, we propose an approach of further exploring the height component by considering that off-ground points of urban scenes collected by a mobile laser scanner (MLS) are usually part of objects which extend in the vertical direction. One of the characteristics of these objects is that they have different heights. For example, a building facade varies in range which may start at eight meter, or higher, depending on the urban area, while the height of a traffic sign mounted on a pole from ground level upwards does usually not exceed 3.0 metres.

2. APPROACH

Our approach is based on assigning height features to each and every point, in addition to exploiting the reflectance value. That means our approach relies on a point-based classification. The two height features we consider within the present research, which is aimed at finding support for our approach, are height difference and number of points within a column spanned up by a cylinder. The cylinder is spanned up around each and every point of the point cloud. The axis of the cylinder is defined by the X, Y coordinates of the point under consideration and the vertical (Figure 1). The first height component used as feature is the height difference Δh which is calculated as the difference of the highest point in the cylinder and the point with the lowest height: $\Delta h = h_{max} - h_{min}$.

The presence of a high number of points within the cylinder indicates an object with a large vertical extension. Therefore, we take as second height component the number of points present in the cylinder. Of course, the height of a column and the number of points in that column are highly correlated. However, in MLS point clouds (and also point clouds acquired

by terrestrial laser scanners) the point density decreases with increasing distance from the sensor. This makes it feasible to explore both point density and height difference simultaneously as features in the classification. So, at the present status of the research we use three features in total: reflectance value (R) given as a digital number (DN) in a range up to 255; number of points (NP) and height difference (Δh). As we do not aim at finding the best classifier but to test the feasibility of exploring height components as features, we use a popular classification scheme which is available as open source software: LIBSVM (Chang and Lin, 2011). The Support Vector Machine (SVM) classifier belongs, together with Random Forest classification (Chehata et al., 2009) and Conventional Neural Network (Fukushima, K. et al., 1982), to the popular classification methods for multiple different application areas, of which the classification of MLS point clouds is just one. Recently the SVM classifier has also been used by Yang et al. (2017) for automatic classification of road facilities from MLS point clouds.
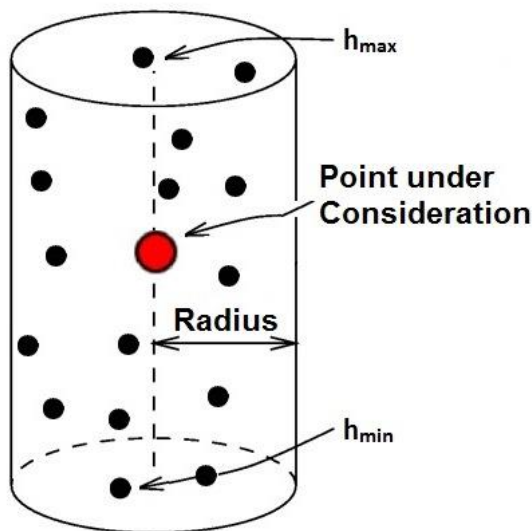

Figure 1. Vertical neighbourhood definition using a cylinder with a pre-specified radius; the height is determined by the lowest and highest point.

3. DATA AND EXPERIMENT

For examining the feasibility of our approach, we conducted experiments using a benchmark dataset generated by the Robotics laboratory (CAOR) at MINES ParisTech, France (Serna et al., 2014). This MLS point cloud, acquired during winter time on February 8, 2013, contains 20 million points and captures a 160m-traject along rue Madame, a street in the 6th Parisian district (France) between rue Madame and rue Vaugirard (Figure 2). Per second 700,000 points were acquired with a precision of 2cm. The observation part of the benchmark dataset consists of the planar coordinates (X, Y) of each point, the height above a reference surface and a DN representing the reflectance value. The developers of the dataset further assigned to each of the 20 million points label and class. In total, 26 classes have been distinguished assigned to the points reflected on 642 individual objects.

A common approach when classifying MLS point clouds is separating ground points from off-ground points. Also, we did

so by leaving the class ground point out of consideration in our experimentation. In sequel research stage we will apply a ground filtering approach. An important tuning parameter of our approach is the radius of the base of the cylinder. After some testing, we selected a fixed radius of 25cm, which is reasonable given the point density. To evaluate our approach not only by means of accuracy measures but also by comparison with other approaches we followed Weinmann et al. (2014) in selecting the classes: facades, cars, pedestrians, motorcycles and traffic signs, with exception of the class ground. A further rationale to choose five classes is based on the fact that the chosen classes are dominantly present within the scene. The points belonging to other classes only count for 0.05% of the complete point cloud. We followed also Weinmann et al. (2014) in the definition of the training set by selecting randomly 1,000 points per class from the total area. To check the distinctiveness of the three features chosen we calculated the mean values and standard deviations of these features in the training set (Table 1). As an example, Figure 3 provides a visual impression of the dispersion of the reflection values of cars. Weinmann et al. (2014) used 26 features mainly derived from exploiting the local geometric structure. This relatively large number of features allowed them to apply Random Forrest classification. In contrast, we used the SVM classifier as said earlier.
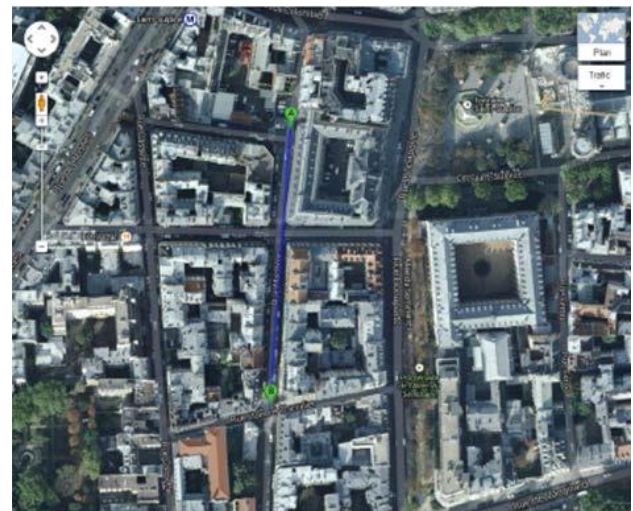

Figure 2. Orthophoto of the benchmark site: Rue Madame, Paris, France (Courtesy: IGN-Google Maps).

		R [DN]	NP	Δh [m]
Façade	μ	153	912	11.7
	σ	44	639	4.5
Cars	μ	206	2869	1.6
	σ	51	1940	2.3
Pedestrians	μ	148	1110	2.7
	σ	51	871	3.3
Motorcycles	μ	215	1399	1.2
	σ	46	828	0.43
Traffic signs	μ	169	2027	2.9
	σ	81	812	0.46

Table 1. Mean (μ) and standard deviation (σ) of the features calculated from the 1,000 training samples per class; R: reflectance value; NP: number of points within cylinder and Δh height difference.
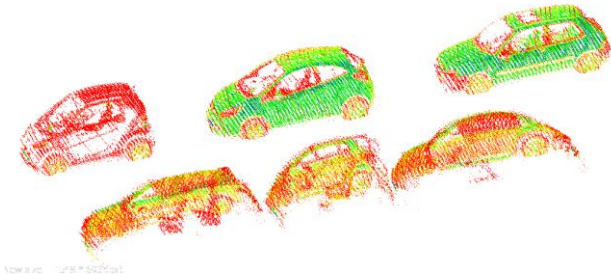
Figure 3. Visualization of the dispersion of the reflectance values of cars (ranging from 55 to 255); the colours range from dark green (lowest) to yellow, red and brown (highest).

4. RESULTS AND DISCUSSION

To evaluate the classification results a confusion matrix was computed by confronting the class assignments to the individual points with the labels of the benchmark dataset. The confusion matrix itself is already a very valuable tool for examining the accuracy (Table 2). Further, compound measures can be derived including the mapping accuracy (MA) per class and overall accuracy (OA). These are also included in Table 2

	Facade	Cars	Pedestrians	Motorcycles	Traffic signs	Sum	MA
Facade	7,744,529	209,660	807,226	672,369	543,651	9,977,435	0.76
Cars	135,033	928,081	209,587	473,755	87,927	1,834,383	0.50
Pedestrians	469	124	3,988	2,129	2,338	9,048	0.44
Motorcycles	218	19,618	25,271	41,690	11,070	97,867	0.43
Traffic signs	2,689	655	2,482	843	7,811	14,480	0.56
Sum	7,882,938	1,158,138	1,048,554	1,190,786	652,797	11,933,213	
						OA	0.73

Table 2. Confusion matrix (i.e. numbers of points) of the five classes showing also the mapping accuracy (MA) and the overall accuracy (OA)

The experiments show that an overall accuracy of 73% could be achieved. Using 26 features Weinmann et al. (2014) achieved an overall accuracy of 90.5%, which is of course significantly higher than our result. Figures 4, 5 and 6 compare the results of both approaches in greater details. The definition of precision, recall and F1-score are given in the Appendix. Façades and cars show high precision. These are also the classes which are very often present in the data set and show significant difference in average reflectance value: 153 versus 206 (see Table 1). 543,651 points of class Facade are assigned to class Traffic signs, which far exceeds the correctly assigned number 7,811 (see Table 2) and results in a low precision. The same is true for the classes pedestrians and traffic signs. To the low precision of the pedestrians contributes the fact that 209,587 points of class Cars are incorrectly assigned to class Pedestrians, only 3,988 points are correctly assigned. The values of the recall measures show a similar trend. As an example, Figure 7 shows the classification of points reflected on cars.

Nevertheless, the results demonstrate that our approach has potential, although it requires a number of significant refinements to obtain accuracies which are high enough for practical application and are at the same level as the results mentioned above. We discuss some of the possible refinements below.

As is the case for many classification pipelines of MLS point clouds also our approach suffers from clutter. The results are also affected in cases where two or more objects are located on top of each other; for example, a pedestrian walking underneath the crown of a tree. If there is a vertical gap in between the objects, they may be separated based on the height of the gap, which requires the setting of a threshold and hence introduces an additional tuning parameter. If there is no vertical gap present, different types of objects may be distinguished on basis of their shape. For example, a tree crown may show significantly different distributions of normal vectors than the traffic sign underneath the tree.

Different objects belonging to the same class may have different sizes. The setting of different radius sizes of the base of the cylinder may support to solve the issues introduced by different object sizes. That means, the setting of different radius sizes may be a means for introducing a multiscale approach and to improve the classification result in this way.

Our approach enables to include statistical measures, particularly standard deviations and covariance matrices, which allows the application of the Maximum Likelihood Classifier (MLC), (see Table 1) a conventional classifier described in many textbooks, e.g. Duda and Hart (1973), and in the remote sensing literature, e.g. Foody et al. (1992). MLC has been thoroughly applied for the classification of multispectral images and also for statistical pattern recognition applications. We think it is worthwhile to conduct, in addition to the popular SVM and Random Forest classifiers and the meanwhile very popular Convolutional Neural Network (CNN) classifier, experiments with the MLS classifier.

Referring to Table 1, two groups can be distinguished based on the reflectance values. Group 1: facades, pedestrians and traffic signs; and group 2: cars and motorcycles. Such a coarse distinction on basis of one feature alone allows for stratified classification.
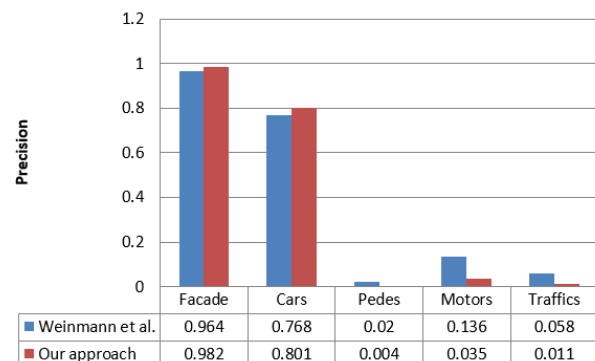

	Facade	Cars	Pedes	Motors	Traffics
Weinmann et al.	0.964	0.768	0.02	0.136	0.058
Our approach	0.982	0.801	0.004	0.035	0.011

Figure 4. Precision values in the five classes

	Facade	Cars	Pedes	Motors	Traffics
■ Weinmann et al.	0.958	0.603	0.559	0.657	0.978
■ Our approach	0.764	0.503	0.442	0.427	0.557
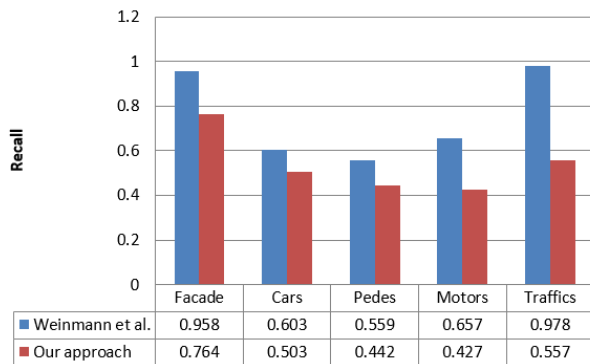
Figure 5. Recall values in the five classes

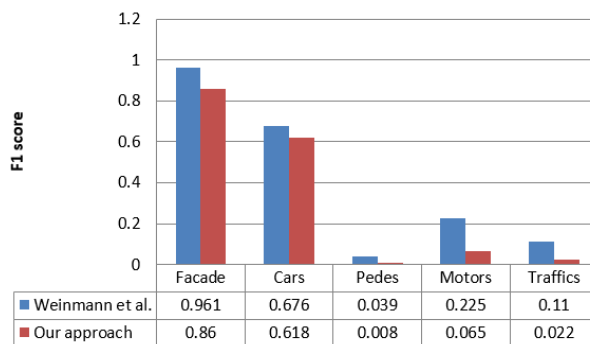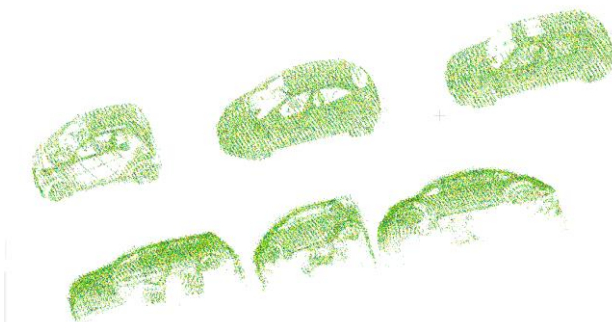|	Facade	Cars	Pedes	Motors	Traffics
■ Weinmann et al.	0.961	0.676	0.039	0.225	0.11
■ Our approach	0.86	0.618	0.008	0.065	0.022

Figure 6. F1-score values in the five classes

Figure 7. Classification of points reflected on cars; cars: green, façades: blue, motorcycles: orange, traffic signs: purple, pedestrians: yellow

5. CONCLUSION AND FUTURE WORK

The fully automated classification of MLS point clouds is still in its infancy. However, the diverse research efforts are rapidly growing. This work hopes to contribute to the development of research directions by recognizing that urban and also road scenes are complex outdoor scenes, which differ a lot from many other scenes which are subject of research in computer vision. In particular, the classification of MLS point clouds is directed towards 3D mapping, that is, the outlining of objects which are of interest for a particular task at hand and the assignment of class labels to these objects (Lemmens, 2017). Our work aims to contribute particularly to automatic class assignment. Although we only used three features in the classification we were able to achieve relatively high classification accuracy. This means that our results are promising and we will continue our research in the direction chosen. In particular we want to focus on including features such as range, i.e. distance between sensor and object and histograms of the height distribution of points within the

cylinder. Furthermore, scene knowledge can be exploited for checking and improving classification results.

Topics for further experimentation are the setting of the radius size of the cylinder, feasibility of using different radius sizes for a multiscale approach, impact of the maximum likelihood classifier on the classification result, selection of the number of training samples and their distribution over the scene, robustness over space and time of training sample selection for use on multiple scenes and use of other benchmark datasets such as the Oakland MLS 3D Point Cloud Dataset of an urban area (Munoz et al., 2009). The efficient use of the sheer amount of points requires a DBMS which may be equipped with classification functionality to avoid data transfer, and so reduce computation time (Cura, et al., 2016). For the latter purpose also spatial indexing and clustering of the points have to be considered. We want also to contribute to the concept of smart point cloud (Poux, et al., 2016).

REFERENCES

Chang, C., Lin, C., 2011. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST), 2(3), pp. 27.

Chehata, N., Guo, L., Mallet, C., 2009. Airborne lidar feature selection for urban classification using random forests. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Paris, France, Vol. XXXVIII, Part 3/W8, pp. 207–212.

Cura, R., Perret, J., Paparoditis, N., 2016. Implicit LOD for processing, visualisation and classification in Point Cloud Servers. arXiv preprint arXiv:1602.06920.

Duda, R.O., Hart, P.E., 1973, Pattern classification and scene analysis, John Wiley & Sons, New York, ISBN 0471223611

Foody, G., Campbell, N., Trodd, N., Wood, T., 1992. Derivation and applications of probabilistic measures of class membership from the maximum-likelihood classification. Photogrammetric engineering and remote sensing, 58(9), pp. 1335-1341.

Fukushima, K., Miyake, S., 1982. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. Competition and cooperation in neural nets. Springer, Berlin, Heidelberg, pp. 267-285.

Grilli, E., Menna, F., Remondino, F., 2017. A review of point clouds segmentation and classification algorithms, Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XLII-2/W3, 2017, 3D Virtual Reconstruction and Visualization of Complex Architectures, 1–3 March 2017, Nafplio, Greece, pp. 339-344.

Guan, H., Li, J., Cao, S., Yu, Y., 2016. Use of mobile LiDAR in road information inventory: a review, International Journal of Image and Data Fusion, 7(3), pp. 219-242.

Lemmens, M., 2017. Mobile Laser Scanning Point Clouds: Automatic Classification and Mapping of Road Objects, GIM International, vol. 31(8).

Munoz, D., Bagnell, J., Vandapel, N., Hebert, M., 2009. Contextual classification with functional max-margin Markov

networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 975–982.

Poux, F., Neuville, R., Hallot, P., Billen, R., 2016. Smart Point Cloud: Definition and Remaining Challenges. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(W1), pp.119-127.

Serna, A., Marcotegui, B., Goulette, F., Deschaud, J., 2014. Paris-rue-Madame database: a 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods. The fourth International Conference on Pattern Recognition, Applications and Methods ICPRAM, Lisbon, Portugal.

Weinmann, M., Jutzi, B., Mallet, C., 2014. Semantic 3D scene interpretation: a framework combining optimal neighborhood size selection with relevant features. Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Zurich, Switzerland, Vol. II-3, 2(3), pp. 181.

Yang, B., Dong, Z., Liu, Y., Liang, F., Wang, Y., 2017. Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, 126, pp. 180–194

Zheng, M., Lemmens, M., van Oosterom, P., 2017. Tentative Tests on Two Rapid Multispectral Classifiers for Classifying Point Clouds. https://agile-online.org/images/conference_2017/ Proceedings2017/posters/126_PosterAbstract_in_PDF.pdf.

APPENDIX

Measure	Description
$F_{i,1,1}$	true positives, point samples in class i are assigned to the class i.
$F_{i,1,2}$	false positives, point samples in class i are assigned to other classes.
$F_{i,2,1}$	false negatives, point samples in other classes are assigned to the class i.
$F_{i,2,2}$	true negatives, point samples in other classes are assigned to other classes.

Table 3. Four common measures.

From the four measures in Table 3 three evaluation measures can be derived.
Precision: a measure of exactness or quality.

$$P_i = \frac{F_{i,1,1}}{F_{i,1,1} + F_{i,2,1}} \qquad (1)$$

Recall: a measure of completeness or quantity.

$$R_i = \frac{F_{i,1,1}}{F_{i,1,1} + F_{i,1,2}} \qquad (2)$$

F1-score: combines precision and recall with equal weights.

$$F1_i = \frac{2P_iR_i}{P_i + R_i} \qquad (3)$$