

INDIVIDUAL ROCKS SEGMENTATION IN TERRESTRIAL LASER SCANNING POINT CLOUD USING ITERATIVE DBSCAN ALGORITHM

A. Walicka¹, G. Józków¹, A. Borkowski¹

¹ Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences, Poland
(agata.walicka, grzegorz.jozkow, andrzej.borkowski)@upwr.edu.pl

Commission II, WG II/10

KEY WORDS: DBSCAN, rocks segmentation, terrestrial laser scanning, fluvial transport

ABSTRACT:

The fluvial transport is an important aspect of hydrological and geomorphologic studies. The knowledge about the movement parameters of different-size fractions is essential in many applications, such as the exploration of the watercourse changes, the calculation of the river bed parameters or the investigation of the frequency and the nature of the weather events. Traditional techniques used for the fluvial transport investigations do not provide any information about the long-term horizontal movement of the rocks. This information can be gained by means of terrestrial laser scanning (TLS). However, this is a complex issue consisting of several stages of data processing. In this study the methodology for individual rocks segmentation from TLS point cloud has been proposed, which is the first step for the semi-automatic algorithm for movement detection of individual rocks. The proposed algorithm is executed in two steps. Firstly, the point cloud is classified as rocks or background using only geometrical information. Secondly, the DBSCAN algorithm is executed iteratively on points classified as rocks until only one stone is detected in each segment. The number of rocks in each segment is determined using principal component analysis (PCA) and simple derivative method for peak detection. As a result, several segments that correspond to individual rocks are formed. Numerical tests were executed on two test samples. The results of the semi-automatic segmentation were compared to results acquired by manual segmentation. The proposed methodology enabled to successfully segment 76% and 72% of rocks in the test sample 1 and test sample 2, respectively.

1. INTRODUCTION

The fluvial transport is an important aspect of hydrological studies. The investigation of different-size fraction movements enables to determine the river bed parameters (such as roughness) and to identify the frequency and the nature of weather events that are necessary to trigger movement of the particular fractions of rock rubble.

Hydrological measurements that are traditionally performed to observe the fluvial transport do not allow determining the displacement of the largest fraction (Józków et al., 2016) or are limited only to the short-term movement detection (e.g. tracers). As a result, various surveying techniques, such as total station measurements, global positioning system, bathymetry, laser scanning, and photogrammetry are used to determine its movement (Picco et al., 2013). The most common method that is applied to investigate and to quantify changes in the river bed is the analysis of DSMs and DSMs of difference (DoDs) (eg. Theule et al., 2012; Barnhart and Crosby, 2013; Picco et al., 2013; Bertin et al., 2015; Kuo et al., 2015). Despite the fact that the DoD method is fast and enables rough estimation of the volume of transported rock material, it does not allow to track the movement of individual stones.

The horizontal movement of individual rocks can be investigated by means of point cloud data analysis. This approach was proposed by Józków et al. (2016) and Lotsari et al. (2015), but in both cases the detection of individual stones was performed manually.

In this work the first step for a semi-automatic rock movement

tracking has been made. As a result, a semi-automatic algorithm for rock segmentation in the TLS point cloud has been proposed.

2. MATERIALS AND METHODS

2.1 Test site and data acquisition

The measurements were performed on Łomniczka river, which is a mountain stream located in the south-western Poland, in Giant Mountains. The study area includes 30 meters long section of the river with local slope of channel over 133 ‰. The Łomniczka river was chosen based on two main reasons. Firstly, the river bed contains massive parts of rocks, which provided suitably large data set. Secondly, during the autumn, the water level in the river is low, which allows to collect TLS points for the lower parts of the rocks.

The terrestrial laser scanning data was collected using Leica ScanStation C10 laser scanner, which is a green-wavelength laser scanner that enables to reach 4 mm range accuracy. The measurements were performed at two scanning positions: first one was located in the beginning of the test site (Figure 1), second one was placed in the middle of the river section. Two point clouds were georeferenced and co-registered using three evenly distributed targets. The global position of each target was measured with centimeter accuracy. Two scanning positions enabled better reproduction of rock shapes and minimization of occlusions.

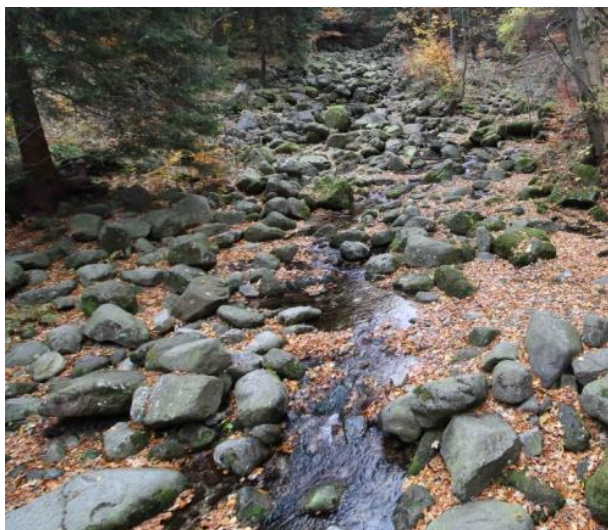


Figure 1. Test site

2.2 DBSCAN algorithm

DBSCAN is a density based clustering algorithm proposed by Ester et al. (1996). The algorithm needs two parameters: the radius of the neighborhood (R), and the number of points located in the defined neighborhood (N). While processing of this algorithm, a point can be labeled as a core, border or noise point. The point p is labeled as:

- core when the number of points in its neighborhood of radius R is bigger than N .
- border when the number of points in its neighborhood of radius R is less than N but the point p is in the neighborhood of another point q .
- noise when the point is not labeled as core or border.

According to (Huang et al., 2017) the algorithm works as follows. All points in the data set are marked as not visited. The algorithm starts from any point that was not visited yet. If the number of points located in the defined neighborhood of point p is not lower than N , the point is added to a new cluster C_i . All the points from the neighborhood of p are assigned to set S and p is marked as visited. Then, all not visited points in the set S are checked. If number of points in the neighborhood of checked point is not lower than N , points from its neighborhood are added to S . If checked point does not belong to any cluster, it is added to C_i . This procedure is repeated for all points from the dataset.

2.3 Peak detection algorithm

As a peak detection method, the simple derivative algorithm was chosen. This method is based on mathematical approach for finding local maxima of function, which relies on finding the point where the first derivative changes its sign. Despite the fact that the simplicity is a big advantage of this method, its main disadvantage is its high sensitivity to the occurrence of noise. Therefore, it was necessary to preprocess the data. The data was smoothed with the mean filter, which window size was set to approximately 5 times more than point spacing. Peak detection algorithm is applied on the data projected into a profile in the principal component domain. This approach is used in order to check, within the rock segmentation process, whether a segment of points represents only one stone or more than one stone.

2.4 Point cloud segmentation

The rocks segmentation is divided into two steps. Firstly, the binary classification is performed which results in point cloud division into two classes: rocks and background. The classification is performed using CANUPO, which is a multi-scale classification algorithm. The multi-scale classification in CANUPO is related to the size of spherical neighborhood of the point. For that reason, CANUPO algorithm requires a parameter that is a set of increased neighborhood sizes. For the purposes of this paper, the classification is based only on geometrical information that can be obtained from the registered point cloud. No radiometric information is included into the classification process. Secondly, the DBSCAN algorithm is iteratively launched on point cloud classified as rocks to divide it into several segments corresponding to individual rocks.

The proposed algorithm works as follows (Figure 2):

- 1) The classification of point cloud is performed using the CANUPO algorithm proposed by Brodu and Lague (2012) and implemented in CloudCompare Software (Girardeau-Montaut, 2016). The goal of this step is to classify the points into two subsets: points representing rocks and background.
- 2) The DBSCAN algorithm is launched on point cloud classified as rocks. In this step the point cloud is divided into small segment.
- 3) For each segment the points are projected on XY plane and the principal component analysis is performed on projected points, which enables to determine the greatest and the lowest variance direction. The first principal component (PC1) corresponds to the greatest variance.

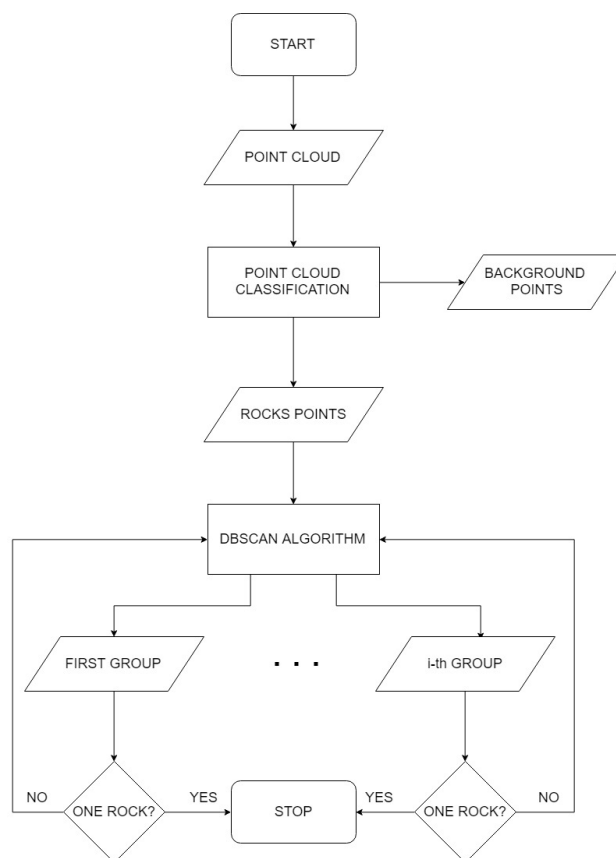


Figure 2. The flow chart of proposed algorithm

4) All points belonging to each segment are projected on PC1Z and PC2Z planes and are divided into equal sections with respect to PC1 and PC2 axes, respectively. For each section the representative value is the maximum value. As a result, 2 profiles showing the representative points are formed.

5) Both profiles prepared in this way are the subject of further analysis based on the detection of peaks representing different stones.

6) If only one peak is found in both profiles, the processing for this segment is finished, and the stone is detected (this segment should contain points belonging to one stone only). If more than one peak is found in any profile, the parameters of DBSCAN are decreased and DBSCAN is executed again, but only on points belonging to this segment. This results in a few new point cloud segments. The processing is performed until all segments contain points belonging to single stone or the N parameter is less than 1.

3. RESULTS

3.1 Data classification

To obtain both the training and validation labels, the manual segmentation of rocks was necessary. Therefore, the training of the classifier and validation of the results were performed on sample data (test sample 1) with the area of 30 m² and mean point density of 11 500 pts/m². The area is characterized by varied structure of the rock fraction and terrain height. It contains both large rock fractions, which will be the subject of further processing and small fractions that should be classified as background points.

For the purposes of the classification, the point cloud was randomly divided into the training and validation sets in proportions of 30% and 70%, respectively. Then the CANUPO algorithm was used to carry out the classification using only geometry information. In order to obtain the optimal results, a series of tests was performed to determine the correct classification parameters. The results of classification tests are compared in Table 1.

Test number	Sizes of the neighborhood [cm]	Kappa score	Overall accuracy
1	5	0.51	0.76
2	10	0.50	0.75
3	15	0.36	0.68
4	20	0.21	0.60
5	[5, 10]	0.57	0.78
6	[10, 15]	0.50	0.75
7	[5, 10, 15]	0.57	0.78
8	[5, 10, 20]	0.57	0.79
9	[30, 50, 70, 90]	0.53	0.77
10	[30, 40, 50, 60, 70, 90]	0.54	0.78
11	[10, 30, 50, 70, 90]	0.68	0.84

Table 1. Classification test results for test sample 1

As a result, the 11th strategy was selected due to its best scores. The results of classification performed on test sample 1 are shown in Figure 3. The visual analysis of the results confirms good quality of classification in the test sample 1. Classification errors occurred only in a few places. In most of such cases, small areas of the river bed were misclassified as rocks. With the exception of one case, the shape of the rocks has been preserved.

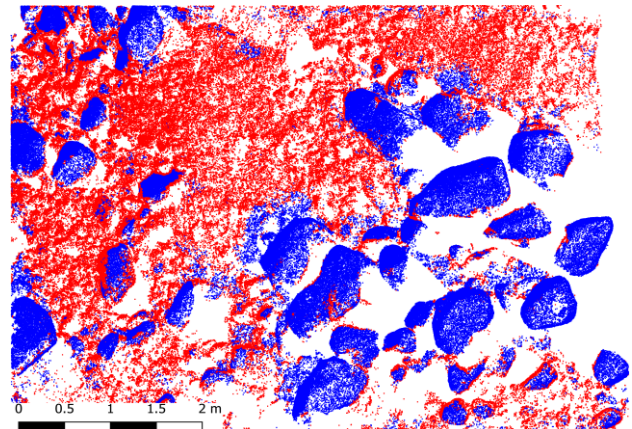


Figure 3. Classification results for the test sample 1
 blue – points classified as rocks
 red – points classified as background

To verify the results of both classification and segmentation, the second test sample was chosen. The test sample 2 has the area of 35 m² and the mean point density of 20000 pts/m². While the first area was characterized by only partially complex rock structures, the second area is mostly covered by ‘overlapping’ rocks of different sizes. The results of classification performed on test sample 2 are shown in the figure 4. In this case, the classification was less accurate than in the case of test sample 1. On this test sample both the large parts of river bed area were misclassified as rocks and parts of stones were misclassified as the river bed.

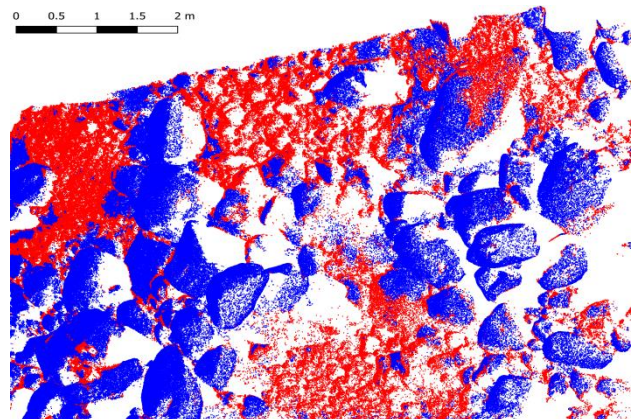


Figure 4. Classification results for the test sample 2
 blue – points classified as rocks
 red – points classified as background

3.2 Point cloud segmentation

The point cloud segmentation was performed on both test samples. The experiments executed on test sample 1 showed that the optimal initial parameters for DBSCAN algorithm, for the processed data, are equal to 5 cm (radius R) and 20 (number of points N located in the defined neighborhood). After each iteration, the N parameter was reduced twice. This enabled the division of larger clusters of rocks into smaller parts representing individual stones or smaller groups of rocks, that could be processed in subsequent iterations (Figure 5).

The results show that 76% of individual rocks in the test sample 1 were segmented correctly. It was necessary to perform 4 iterations. The detailed results of each iteration are compared in Table 2.

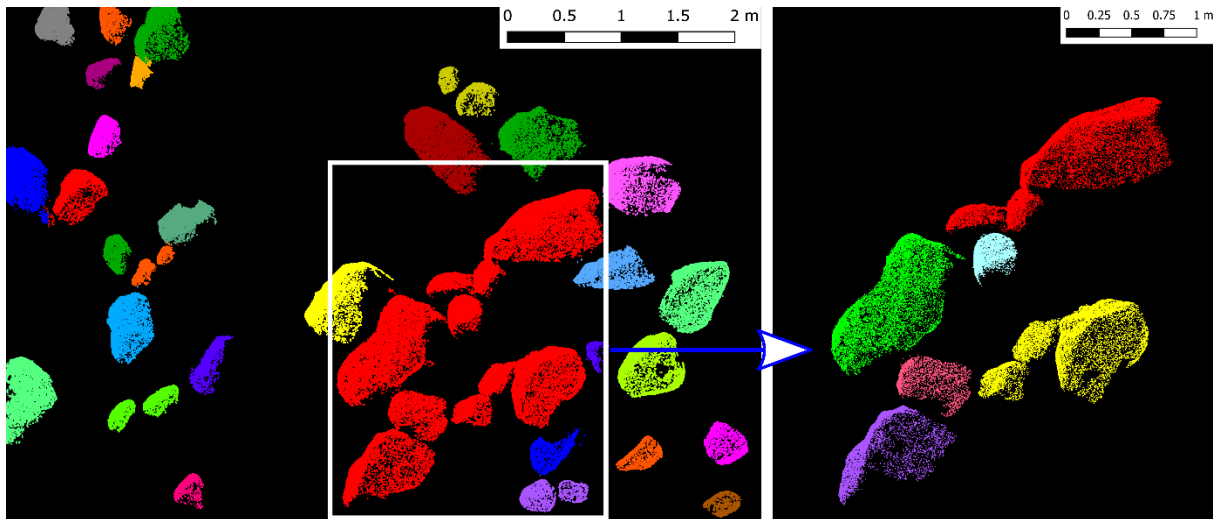


Figure 5. Two first iterations of the proposed algorithm

Iteration	Radius (R)	Number of points (N)	Percent of detected stones
1	5 cm	20	61%
2	5 cm	10	70%
3	5 cm	5	72%
4	5 cm	2	76%

Table 2. Rocks segmentation results for the test sample 1

The similar results were obtained for the test sample 2. The detailed results of each iteration are compared in Table 3. In this case it was possible to segment 72% of individual rocks and 4 iterations were needed. However, on this test sample, data processing was much more demanding than on test sample 1. Due to the bigger complexity of the terrain and worse classification results, in the first iteration it was possible to segment only 18% of rocks. Additionally, the misclassifications of the river bed as rocks caused that, after the DBSCAN iteration, several segments containing only points that belong to the river bed were formed.

Iteration	Radius (R)	Number of points (N)	Percent of detected stones
1	5 cm	20	18%
2	5 cm	10	46%
3	5 cm	5	65%
4	5 cm	2	72%

Table 3. Rocks segmentation results for the test sample 2

CONCLUSIONS

In this study the preliminary results of single stones segmentation in a TLS point cloud are presented. For this purpose the two-step methodology is proposed. Firstly, the point cloud is classified into two subsets containing rocks and background, secondly the rock subset is divided into several segments corresponding to individual stones using DBSCAN algorithm. The DBSCAN algorithm is executed iteratively until only one rock is detected in every segment. The number of stones in each segment is calculated using principal component analysis and simple derivative method for peak detection.

This methodology enabled to segment 76% and 72% of individual rocks present in the test sample 1 and test sample 2, respectively. The essential step that affects the correctness of the algorithm operation is point cloud data classification. The classification errors may lead to the loss of the shape of individual stones and the formation of segments containing only points belonging to the river bed.

In this study the first step of the semi-automatic algorithm for calculating the translation and rotation movement parameters for each individual rock in the mountain river bed has been made. However, the presented results are the effect of the initial studies. The future experiments will aim to increase the accuracy of classification and to eliminate the problem of the formation of segments containing only points belonging to the river bed. Furthermore, due to the character of DBSCAN algorithm, it is clear that its initial parameters depend on point cloud density. Thus, the future experiments will also include both the automatic determination of DBSCAN initial parameters and the investigation on the influence of point cloud density to initial parameters of the algorithm.

REFERENCES

- Barnhart, T.B., Crosby, B.T., 2013. Comparing Two Methods of Surface Change Detection on an Evolving Thermokarst Using High-Temporal-Frequency Terrestrial Laser Scanning, Selawik River Alaska. *Remote Sensing*, 5, pp. 2813-2837.
- Bertin, S., Friedrich, H., Delmas, P., Chan, E., Gimel'farb, G., 2015. Digital stereo photogrammetry for grain-scale monitoring of fluvial surfaces: Error evaluation and workflow optimization. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101, pp. 193–208.
- 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, pp. 121-134.
- Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot-Deseilligny, M., Puech, W., 2007. Processing full-waveform lidar data: modelling raw signals. *International Archives of*

the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 102-107.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X.W. A density-based algorithm for discovering clusters clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996, pp. 226–231.

Girardeau-Montaut, D. 2016. CloudCompare - Open Source project. Online: <http://www.danielgm.net/cc/>

Huang, F., Zhu, Q., Zhou, J., Tao, J., Zhou, X., Jin, D., Tan, X., Wang, L., 2017. Research on the Parallelization of the DBSCAN Clustering Algorithm for Spatial Data Mining Based on the Spark Platform. *Remote Sensing*, 9(12), 1301.

Józków, G., Borkowski, A., Kasprzak, M., 2016. Monitoring of fluvial transport in the mountain river bed using Terrestrial Laser Scanning. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41, pp. 523-528, DOI: 10.5194/isprs-archives-XLI-B7-523-2016.

Kuo, Ch-W., Brierley, G., Chang, Y-H., 2015. Monitoring channel responses to flood events of low to moderate magnitudes in a bedrock-dominated river using morphological budgeting by terrestrial laser scanning. *Geomorphology*, 235, pp. 1–14.

Lotsari, E., Wang, Y., Kaartinen, H., Jaakkola, A., Kukko, A., Vaaja, M., Hyyppä, H., Hyyppä, J., Alho, P., 2015. Gravel transport by ice in a subarctic river from accurate laser scanning. *Geomorphology*, 246, pp. 113-122.

Picco, L., Mao, L., Cavalli, M., Buzzi, E., Rainato, R., Lenzi, M.A., 2013. Evaluating short-term morphological changes in a gravel-bed braided river using terrestrial laser scanner. *Geomorphology*, 201, 323–334.

Weinmann, M., Weinmann, M., Mallet, C., Brédif, M., 2017. A Classification-segmentation framework for the detection of individual trees in dense MMS point cloud data acquired in urban areas. *Remote Sensing*, 9(3), 277.