CIVIL ENGINEERING JOURNAL

Article no. 5

THE CIVIL ENGINEERING JOURNAL 1-2025

POINT CLOUD LOCAL NEIGHBORHOOD FEATURES USED FOR CLASSIFICATION - A REVIEW

Martin Boušek, Jakub Kučera and Hana Váchová

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Special Geodesy, Prague, Thákurova 7, 166 29 Prague 6, Czech Republic; martin.bousek@fsv.cvut.cz

Received: 2.12.2024

Received in revised form: 20.01.2025

Accepted: 14.02.2025

ABSTRACT

Point clouds are essential for 3D spatial analysis and widely used in geodesy, photogrammetry, and remote sensing. While modern technologies simplify their collection, processing remains challenging due to data size, irregularity, and noise. Classification is critical for object identification and noise removal.

This paper explores geometric features of points derived from their local 3D neighbourhoods. It examines neighbourhood definitions, feature computation via principal component analysis (PCA), and their impact on real dataset classification. Using a test point cloud with natural and anthropogenic features, we analyze feature dependencies, identify redundancies, and highlight key metrics. Additionally, we propose new approaches for noise filtering, contributing to more efficient point cloud processing and practical applications.

KEYWORDS

Classification, Filtering, Geometric features, Neighbourhood, Point cloud

INTRODUCTION

Nowadays, non-selective spatial data collection methods are increasingly used to obtain spatial information. The resulting products of these methods are so-called point clouds, which serve as a fundamental representation of 3D space and have a wide range of applications [1]. They are utilized in various fields such as geodesy, remote sensing, photogrammetry, robotics, computer vision, and others [2, 3]. Point clouds are a key source of information in modern applications, making their accurate interpretation essential.

Although modern measurement methods and advanced technologies enable fast and convenient data collection, they also introduce new challenges in post-processing. The size, complexity, and irregularity of the collected data require advanced methods for analysis and information extraction.

To extract meaningful information from a point cloud containing diverse objects, classification is often necessary. This involves assigning a specific value (label) to each point. For example, to separate vegetation from other categories, we can assign a label (e.g., a value of 1) to vegetation points and a different label to other classes. This classification can then be extended into a filtering process, where points with specific labels are removed to clean the point cloud. Significant progress in point cloud classification methods, particularly those using machine learning, has been made in the last decade.

Each point of the point cloud carries with it some information either directly measured or obtained from the data measured by the sensor (XYZ coordinates, RGB color, intensity, echo). We classify this information as a basic measured feature that gives us information about the point itself. Furthermore, from the neighbouring points of each point, it is possible to infer geometric features of the nearby neighbourhood, which can be interpreted as a description of the scanned object.

The paper by Qi et al. [4] proposes a neural network model working exclusively with the basic spatial coordinates of the XYZ point cloud, using which the network segments the cloud into semantic objects. In [5], Štroner et al. use the RGB point cloud information to derive various vegetation indices, using which they filter vegetation from the point cloud. Atik and Duran [6] use RGB information both alone and together with derived geometric information from neighbouring points to classify different types of objects - terrain, buildings, high and low vegetation. Authors Liu et al. [7] classify vegetation and ground points in lidar aerial point cloud using intensity, elevation, scan angle and distance from scanner. They also use features calculated from the point's neighbourhood - eigenvalues, normal vector and geometric features calculated from eigenvalues such as planarity, linearity and eigenentropy. Chehata et al. in [8] also add to these features information about the order of reflection (echo) of the lidar beam.

Other studies such as in Weinmann et al. [9, 10] use only the features of the local 3D neighbourhood computed from eigenvalues to classify and segment the point cloud. They also discuss the relevance of the individual features and the importance of the correct selection of the 3D local neighbourhood. In this approach, the type of data collection does not matter, as it does not use sensor-specific information - RGB color in the case of photogrammetry or reflection intensity and echo in the case of laser scanning - and is thus generally applicable. Therefore, in this paper we will focus exclusively on features computed from the 3D neighbourhood of a point.

However, many of these methods have been proposed and tested on idealized datasets that do not reflect real-world conditions. Real-world data often contains noise, exhibit non-uniform point densities, and are influenced by measurement conditions or technical limitations of the instruments. To ensure correct interpretation and efficient processing, these aspects must be considered, and individual points carefully analyzed based on their local neighbourhood. Therefore, this paper addresses two key questions: 1) what features can be computed for each 3D point in the point cloud, and 2) what object properties in the point cloud are represented by these features. We further analyze feature correlations and propose potential new approaches for their computation.

MATERIALS AND METHODS

Test point cloud

In the point cloud selected for testing (Figure 1), a portion of a railway bridge over a forested valley is captured. The scene was chosen for its diversity, as it contains both anthropogenic features such as straight walls, arches and small engineering features, as well as a variety of natural terrain including trees and slopes. The cloud contains approximately 830,000 points and covers an area of approximately 1,180 m2. The cloud was created by combining data acquired with a Leica P40 laser scanner and data acquired by aerial photogrammetry using a DJI Phantom 4 RTK UAV.

Inconsistencies in the distribution of points in the point cloud may arise due to the combination of the two data collection methods. In order to homogenize the data, the point cloud was subsampled to 7 cm and since we used a sufficiently large neighborhood when calculating the characteristics, the differences in methods will not be apparent. The two methods we used are compared in terms of accuracy in the papers [11, 12], where the authors determined a difference in accuracy of the methods in the order of centimeters. For the purposes of our testing, such accuracy between point clouds from different methods is sufficient.

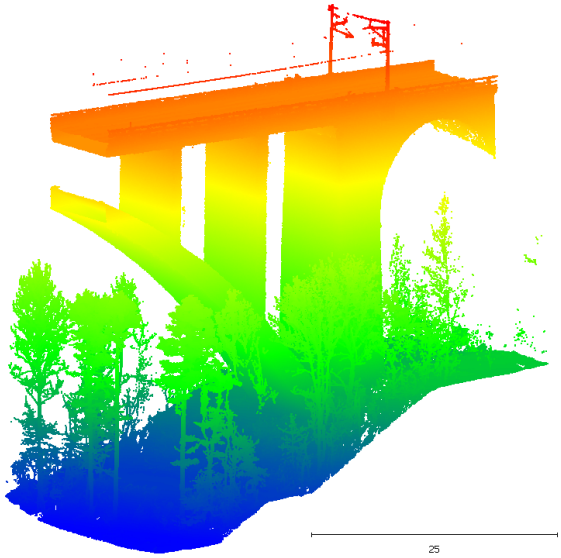

Fig. 1 – Test point cloud colored based on elevation

GEOMETRIC FEATURES

3D neighbourhood

Geometric features are used to describe the relationships between neighbouring points in a point cloud. Therefore, in order to analyze the local 3D structure around a point X using geometric features, it is important to first define the relevant neighbourhood correctly. There are several strategies to define this neighbourhood of a point X: based on a fixed number of nearest neighbours, using a spherical neighbourhood or a cylindrical neighbourhood [8]. The actual neighbourhood is then the set of points falling within this region, i.e. the "neighbours" of point X, which is shown in Figure 2.

The choice of the neighbourhood and the extraction of the features interact because the distinction of the geometric features depends on the relevant neighbourhood into which the 3D points used for feature extraction fall. Some features may be more pronounced in a wider neighbourhood, while others will be more evident in a smaller neighbourhood. The choice of neighbourhood size also depends on the density of the point cloud and its local variations [13].

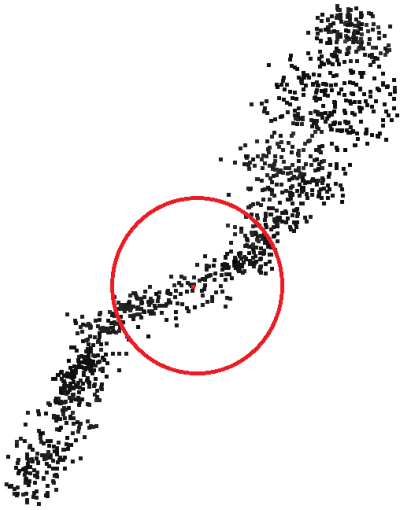

Fig. 2 – 2D representation of a spherical neighbourhood of a point in a point cloud

Calculation of local neighbourhood features

Principal component analysis (PCA) is used to calculate the local features of neighbours of a point in the point cloud. This is a set of unit vectors that give the direction of the lines that best represent a collection of points. Each i-th vector must be perpendicular to all previous vectors. The first one is in the direction of the highest variance of the collection. After the influence of the first vector is removed, the next vector is again found as the direction of greatest variance. The set of such vectors is orthonormal and points to the directions of the axes of the largest variances of the collection, which are linearly uncorrelated with each other. The resulting PCA vectors of the data set are identical to the eigenvectors of its covariance matrix. Its eigenvalues $\lambda_1, \lambda_2, \lambda_3$ then indicate the variance in the direction of these vectors (Figure 3) [14, 15]. For these eigenvalues, it is always the case that $\lambda_1 > \lambda_2 > \lambda_3$. In practice, singular value decomposition (SVD) is used to compute the eigenvalues from the covariance matrix [14].

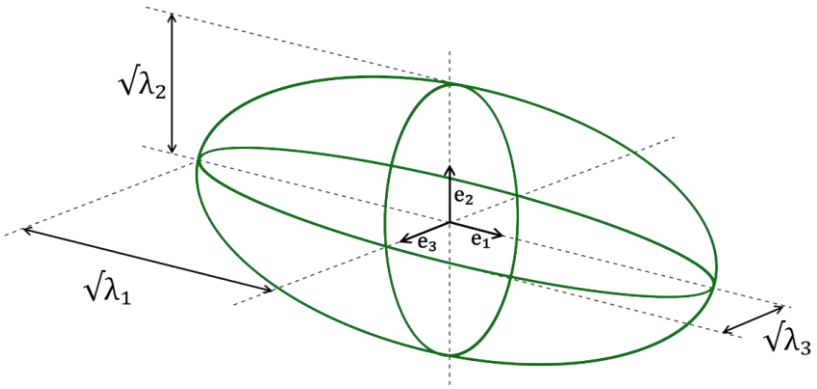

Fig. 3 – A general ellipsoid characterizing the shape of the point's neighbourhood with a representation of the eigenvalues and eigenvectors

Features based on eigenvalues and eigenvectors

Based on this description of a point's neighbourhood, also known as the 3D structure tensor, the derived neighbourhood features can be calculated.

The program for loading point clouds and calculating characteristics was developed in Python using the CloudComPy library. For the calculations, a neighbourhood of 0.5 m was chosen. The examined features are as follows:

Planarity

The value of planarity in the local neighbourhood can be calculated using the Equation 1:

$$P = \frac{\lambda_2 - \lambda_3}{\lambda_1}. \quad (1)$$

This feature describes how closely the distribution of points in the local neighbourhood approximates a plane, as shown in Figure 4 a). Red indicates the highest planarity values (e.g., bridge structure walls), while blue and green shades represent lower values (e.g., vegetation). This visual distinction makes it possible to differentiate planar surfaces from more complex, less planar structures.

Linearity

The linearity in the chosen neighbourhood of the point can be calculated according to Equation 2:

$$L = \frac{\lambda_1 - \lambda_2}{\lambda_1}. \quad (2)$$

This feature determines how much the surrounding points are aligned along a straight line. When λ_2 and λ_3 are small, it indicates that the points are stretched significantly in one direction. Elements with high linearity are highlighted in red in Figure 4 b). This is the most noticeable on the tree trunks and the technical equipment at the top of the bridge.

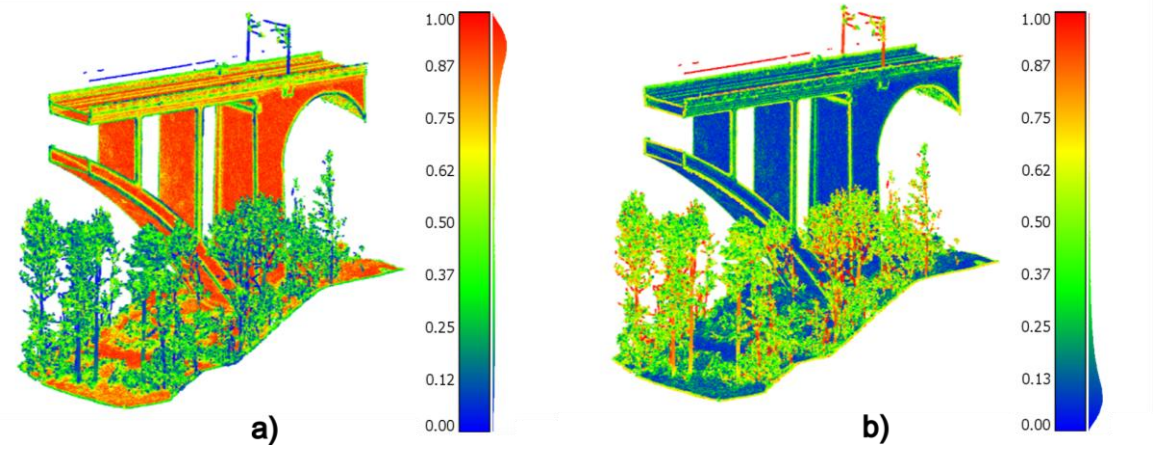

Fig. 4 – Test point cloud colored based on: a) Planarity, b) Linearity

Sphericity

The formula for calculating the sphericity in the local neighbourhood of a point is Equation 3:

$$S = \frac{\lambda_3}{\lambda_1}. \quad (3)$$

Sphericity describes the degree to which the points in the neighbourhood are distributed in all directions. In Figure 5 a), blue represents areas with lower sphericity, while green indicates higher values. The bridge structure appears mostly in dark blue, as expected, because it consists of linear and planar surfaces where the points are not uniformly distributed in all directions but are concentrated in certain directions. In contrast, trees and vegetation show more directionally dispersed points, as the leaves and branches create an irregular, three-dimensional structure. However, the boundary between these two types of features is not entirely clear.

Anisotropy

The value of anisotropy in the local neighbourhood can be calculated using Equation 4:

$$A = \frac{\lambda_1 - \lambda_3}{\lambda_1}. \quad (4)$$

This feature indicates how much the points are arranged along a dominant direction (high anisotropy) compared to a uniform distribution in all directions (low anisotropy). High anisotropy values are present where the points are more linearly distributed (e.g., along one direction). In Figure 5 b), high anisotropy values are visualized in red, highlighting regions with a distinct directional arrangement of points, which typically correspond to engineering structures. In contrast, vegetation exhibits variable anisotropy, reflecting its more complex geometry.

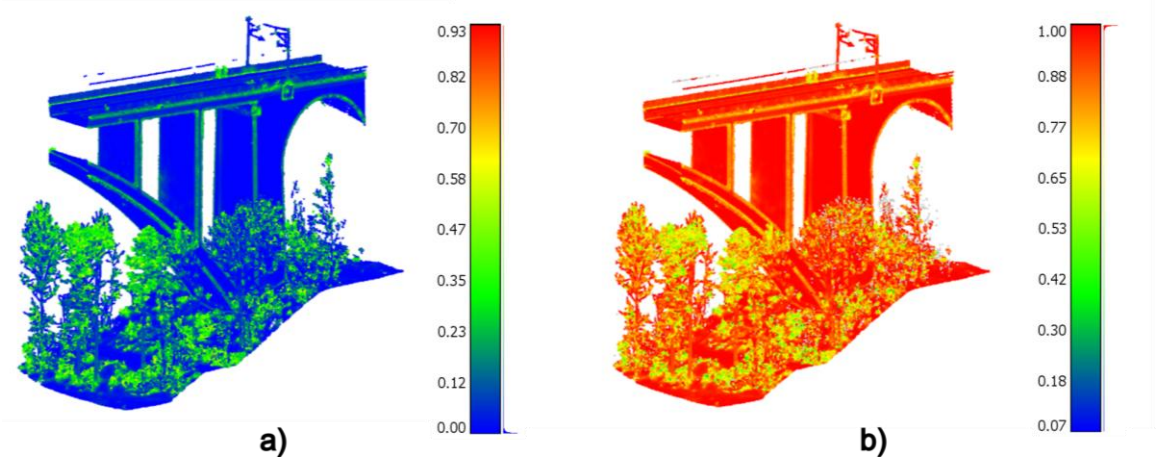

Fig. 5 – Test point cloud colored based on: a) Sphericity, b) Anisotropy

Omnivariance

We can calculate the omnivariance using Equation 5:

$$O = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3}. \quad (5)$$

This feature shows the distribution of points in different directions and is shown in Figure 6 a). When λ_1, λ_2, and λ_3 are similarly large, the omnivariance is higher, indicating a uniform distribution of points. Omnivariance captures the size and scale of the point distribution, relying solely on the magnitude of the eigenvalues.

Eigenentropy

The formula for calculating eigenentropy in a local neighbourhood of a point is Equation 6:

$$E = -\Sigma_{i=1}^{3} \lambda_i \ln \lambda_i. \quad (6)$$

Eigenentropy indicates the degree of "chaos" or uncertainty in the distribution of points. A higher value suggests a more uniform distribution (Figure 6 b)). Unlike omnivariance, eigenentropy is scale-independent and focuses on the relative relationships between the eigenvalues, rather than their absolute magnitudes.

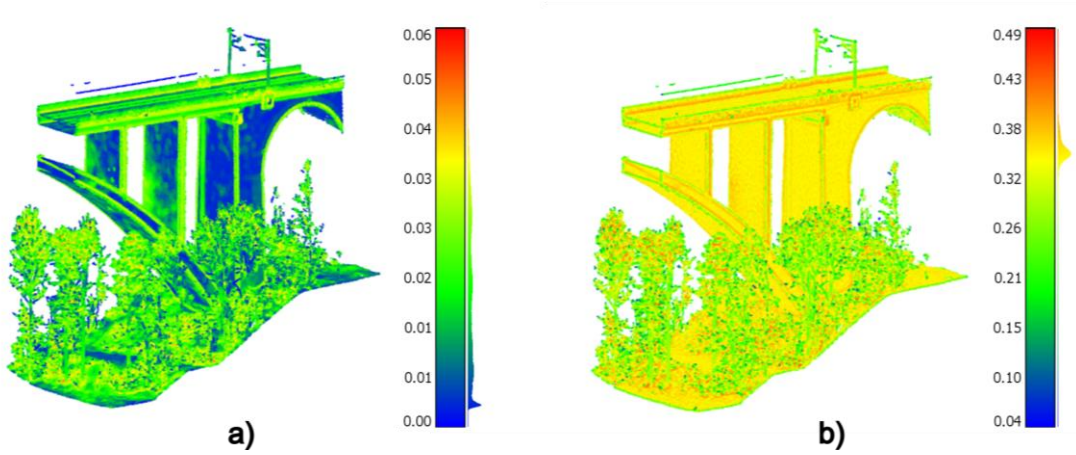

Fig. 6 – Test point cloud colored based on: a) Omnivariance, b) Eigenentropy

Surface variation

Surface variation can be calculated according to Equation 7:

$$Sv = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}. \quad (7)$$

Surface variation describes the deviation of neighbouring points from the plane in a given local neighbourhood, indicating how smooth or rough a surface is. Values close to zero mean the points lie nearly in a plane, while higher values indicate a rough or irregular surface. In Figure 7 a), the main structure of the bridge is predominantly blue, suggesting that the bridge surface is relatively flat, which aligns with expectations for an engineering structure. The figure also implies a potential correlation between this feature and sphericity or anisotropy.

Verticality

The verticality value is calculated by subtracting the absolute value of the scalar product of the unit vector in the direction of the vertical axis and the third eigenvalue λ_3 from one (Equation 8):

$$V = 1 - |\langle [0 \;\; 0 \;\; 1], \lambda_3 \rangle|. \quad (8)$$

Verticality ranges from 0 to 1, with a perfect vertical orientation corresponding to a value of 1. In Figure 7 b), high values are predominantly seen in the vertically oriented bridge walls and the trunks of larger trees. Conversely, low values are found mainly in points located on the horizontal surface beneath the bridge and in points forming the bridge deck.

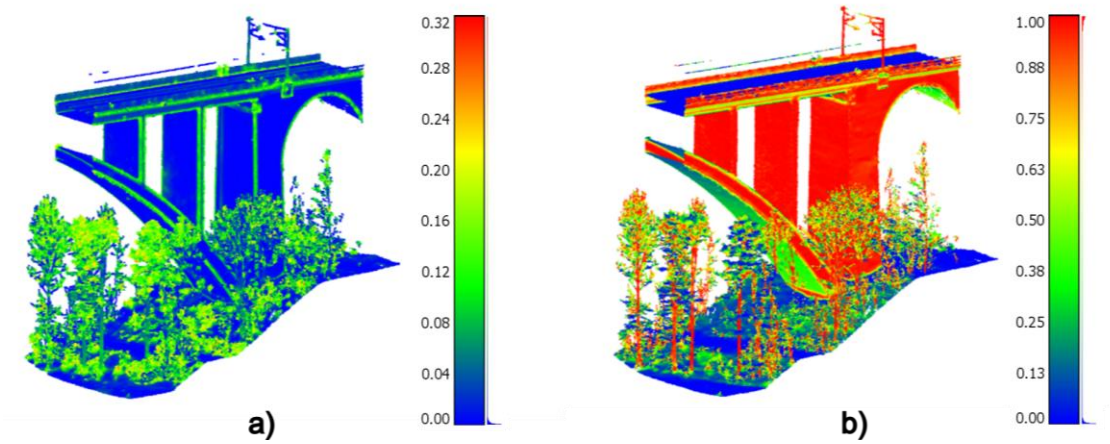

Fig. 7 – Test point cloud colored based on: a) Surface variation, b) Verticality

PCA1

PCA1 equals to a division of the first eigenvalue λ_1 and sum of all eigenvalues (Equation 9):

$$PCA1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3}. \quad (9)$$

A high value (max. 1) indicates that the surrounding points are concentrated in one direction, corresponding to linear structures (e.g., trunks, wires). In this case, the wires above the bridge deck and some tree branches are particularly noticeable in the point cloud. Conversely, low values suggest that the surrounding points are not primarily aligned in one direction, which can be seen in Figure 8 a).

PCA2

This feature is very similar to PCA1, the difference being that the dividend in the equation is λ_2 (Equation 10):

$$PCA2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}. \quad (10)$$

Values close to 0.5 (the maximum value) indicate a planar distribution of points in the selected neighbourhood. As with PCA1, low values suggest the opposite of high values, meaning either a non-planar or omnidirectional distribution of points (Figure 8 b)).

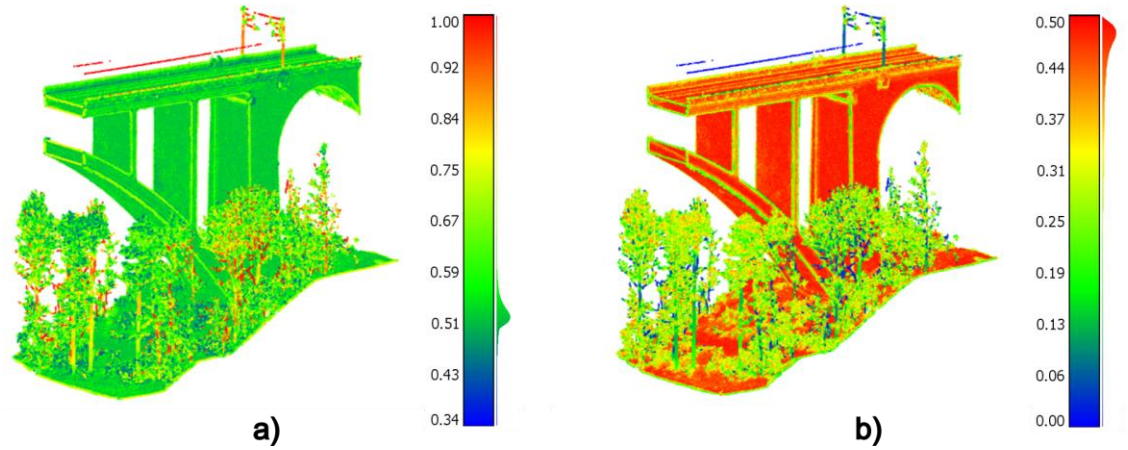

Fig. 8 – Test point cloud colored based on: a) PCA1, b) PCA2

Sum of eigenvalues

The sum of eigenvalues in the selected local neighbourhood can be calculated using Equation 11:

$$\Sigma\lambda_i = \lambda_1 + \lambda_2 + \lambda_3. \quad (11)$$

This feature reflects 3D variance of the neighbouring points. It is the quadratic sum of the standard deviations in the local 3D neighbourhood. The distinct color boundary between object types is not clearly defined in Figure 9, where all features are represented in yellow and green.

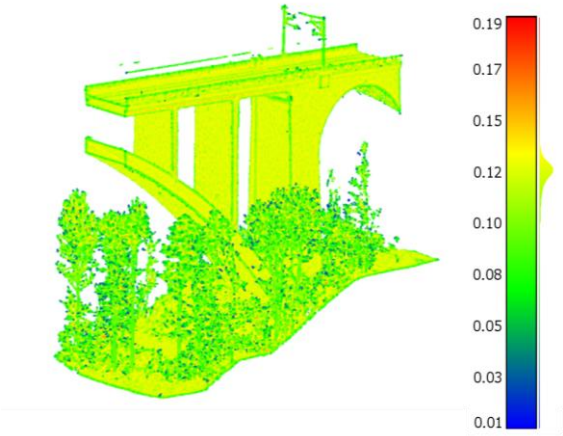

Fig. 9 – Test point cloud colored based on Sum of eigenvalues

Other features

Number of neighbours

The number of neighbours specifies the total number of points, n, in the selected neighbourhood of a given point (Figure 10 a)). As mentioned earlier in the chapter, the choice of neighbourhood size and type is crucial for the eigenvalue calculation. In this case, a spherical neighbourhood with a radius of 0.5 meters was selected to calculate all the presented features, including the number of neighbours.

Distance to plane

Distance to plane expresses distance of point X from plane p and can be calculated using Equation 12:

$$\vartheta(X,p) = \frac{|ax_i + by_i + cz_i + d|}{\sqrt{a^2 + b^2 + c^2}}, \tag{12}$$

where a, b, c, d are the parameters of the point-normal equation of the plane, and x_i, y_i, z_i represent the coordinates of the neighbouring points. The parameters a, b and c correspond to the values of the normal vector calculated through Singular Value Decomposition (SVD), and the parameter d is computed using a, b and c along with the coordinates of the center of gravity through which the plane passes. This feature can be helpful, for example, in identifying noise or low vegetation in point clouds (Figure 10 b)).

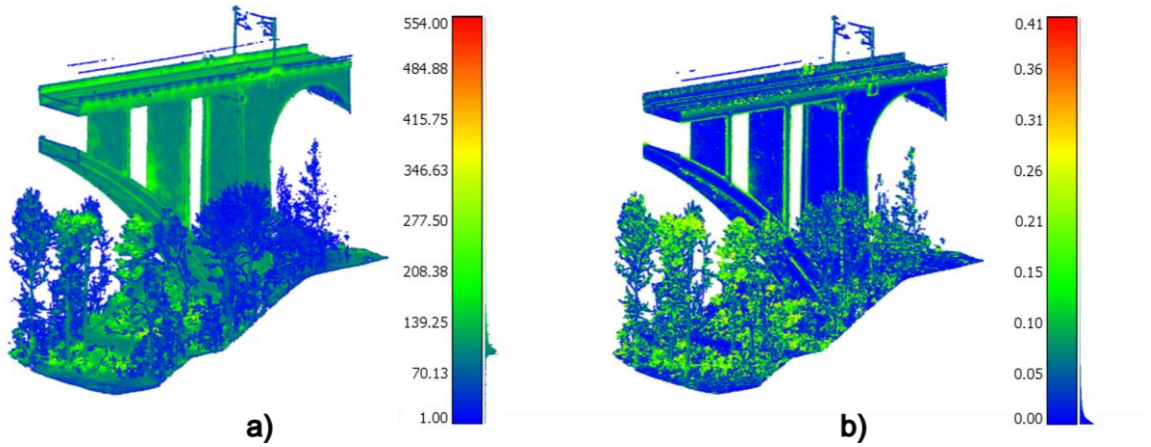

Fig. 10 – Test point cloud colored based on: a) Number of neighbours, b) Distance to plane

Surface density

Surface density expresses the density of points in the surface of a circle of the selected radius. It can be used to infer the density of points per unit area. In a relative color scale, the feature appears to be strongly similar to number of neighbours, as can be seen in Figure 11 a). The value of this feature can be calculated according to Equation 13:

$$Sd = \frac{n}{\pi r^2}, \tag{13}$$

where r is the radius of the local neihbourhood and n is the number of points in the neighbourhood.

Volume density

Volume density expresses the point density within the volume of the selected neighbourhood's sphere. It can be used to infer the number of points per unit volume. In a relative color scale, this feature appears similar to the number of neighbours and surface density features (Figure 11 b)). The value of this feature can be calculated using Equation 14:

$$Vd = \frac{n}{\frac{4}{3}\pi r^3}, \tag{14}$$

where r is the radius of the local neihbourhood and n is the number of points in the neighbourhood.

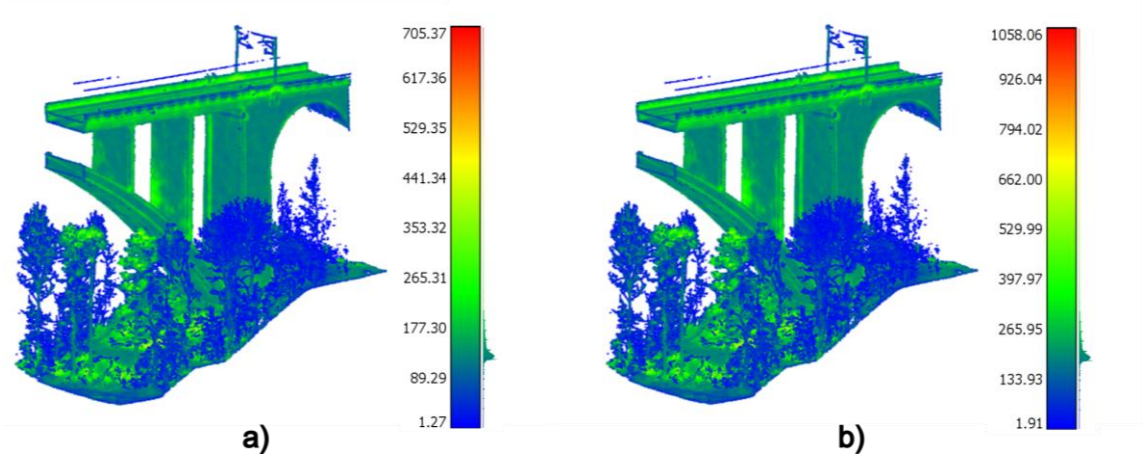

Fig. 11 – Test point cloud colored based on: a) Surface density, b) Volume density

Height standard deviation

This feature shows the standard deviation of the heights of all points in the local neighbourhood. As seen in the Figure 12 a), it can be used to distinguish between horizontal, vertical, and inclined surfaces. The formula for the calculation is described in the Equation 15:

$$s_Z = \sqrt{\frac{\sum_{i=1}^{n}(\bar{Z} - Z_i)^2}{n}}, \tag{15}$$

where n is the number of points in the neighbourhood, Z_i is height of a point of the neighbourhood and $\bar{Z}$ is the mean of the heights of all the points in the neighbourhood.

Max height difference

The feature is calculated as the difference between the height of the highest and lowest point of the neighbourhood. The result is similar to the Height standard deviation and can be used to detect horizontal surfaces (Figure 12 b)).

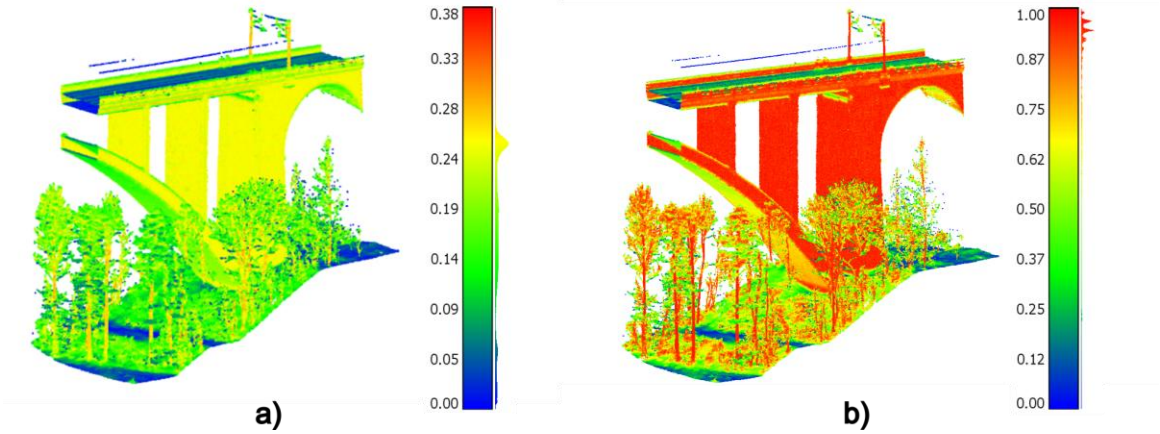

Fig. 12 – Test point cloud colored based on: a) Height standard deviation, b) Max height difference

FEATURE ANALYSIS

The relationships between the individual features describing the geometric and spatial properties of the points in the point cloud were analyzed based on the coefficient of determination. This reflects how well the values of one feature can be explained by the other. The results show the presence of strong, medium and weak dependencies, see Figure 13.

r = 0.5 m	PLA	LIN	SPHE	NON	SOE	OMNI	ENTR	ANI	SV	VD	SD	MHD	DTP	HSD	VERT	PCA1	PCA2
planarity (PLA)																	
linearity (LIN)	0.75																
sphericity (SPHE)	0.51	0.07												degree of determination			
number of neighbours (NON)	0.00	0.07	0.23											0.00-0.69			
sum of eigenvalues (SOE)	0.28	0.41	0.02	0.15										0.70-0.80			
omnivariance (OMNI)	0.38	0.07	0.67	0.31	0.00									0.81-0.90			
eigenetropy (ENTR)	0.06	0.27	0.06	0.35	0.82	0.18								0.91-1.00			
anisotropy (ANI)	0.51	0.07	1.00	0.23	0.02	0.67	0.06										
surface variation (SV)	0.62	0.14	0.97	0.18	0.04	0.70	0.04	0.97									
volume density (VD)	0.00	0.07	0.23	1.00	0.15	0.31	0.35	0.23	0.18								
surface density (SD)	0.00	0.07	0.23	1.00	0.15	0.31	0.35	0.23	0.18	1.00							
max height difference	0.04	0.01	0.07	0.04	0.01	0.06	0.03	0.07	0.08	0.04	0.04						
distance to plane (DTP)	0.00	0.00	0.01	0.01	0.05	0.01	0.05	0.01	0.00	0.01	0.01	0.01					
height standard deviation (HSD)	0.02	0.01	0.02	0.01	0.00	0.02	0.01	0.02	0.03	0.01	0.01	0.94	0.00				
verticatlity (VERT)	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.66	0.00	0.77			
PCA1	0.19	0.66	0.09	0.26	0.30	0.06	0.46	0.09	0.04	0.26	0.26	0.01	0.01	0.00	0.01		
PCA2	0.91	0.92	0.24	0.01	0.37	0.19	0.18	0.24	0.34	0.01	0.01	0.02	0.00	0.02	0.00	0.46	

Fig. 13 – Coefficients of determination of the examined features

The analysis revealed a complete dependence between anisotropy and sphericity. This result suggests that both features share common geometric properties of the point cloud, and their calculation is strongly linked. One hundred percent dependence was also found between the features number of neighbours, volume density and surface density.

Strong coefficient of determination was found between anisotropy and surface variation (0.97), sphericity and surface variation (0.97) or between maximum height difference and height standard deviation (0.94).

From looking at the point cloud and the distribution of features in it, it was concluded that some of them do not have significant predictive value for filtering and classifying point clouds. For example, we have identified omnivariance, sum of eigenvalues, eigenentropy and PCA1 as such. The features planarity, surface variation, PCA2 and distance to plane may be suitable for classification and filtering. Some features, which work for example with the height distribution of points in the neighbourhood, are mainly suitable for creating digital terrain models and for finding horizontal surfaces. These features include verticality, height standard deviation, max height difference.

This analysis suggests that some features of the point neighbourhood are redundant, while others may be of great importance for describing the point cloud.

CHALLENGES AND FUTURE DIRECTIONS

Currently used features calculated from the local 3D neighbourhood of a point work well in the case of resolution of e.g. flat surfaces and complex objects (vegetation), where the classification is reliable. The description of a set of points by eigenvectors is in fact a fitting of the points of the neighbourhood with a plane. However, in the real world, flat surfaces do not often occur and when we change the scale (the size of the local 3D neighbourhood of a point) we can get more complex surfaces. Because of this, the difference between the features of each object can then be unclear and indistinguishable to an automated program. Therefore, when using large neighborhoods, we suggest fitting the neighbourhood points with a higher order surface and determine the deviations of individual points from this surface in the direction of the 3rd eigenvector. Such a surface can better represent the studied object, and for example, noise removal can then be more effective. A possible limitation of this approach in dense point clouds may be the high demand on computer power when calculating the surface using the least squares method. It is used for adjustment of the second order surface parameters according to Equation 16:

$$z_{r,i} = a_0 + a_1 x_{r,i} + a_2 y_{r,i} + a_3 {x_{r,i}}^2 + a_4 {y_{r,i}}^2 + a_5 x_{r,i} y_{r,i}, \tag{16}$$

where $a_0 - a_5$ are the second order surface parameters and $x_{r,i}, y_{r,i}, z_{r,i}$ are the coordinates of the i-th point in the local neighbourhood, which are rotated for calculation purposes so that their third eigenvector is identical to the direction of the Z-axis. After alignment, the deviations of the points from the surface in the direction of the third eigenvector v are calculated and the standard deviation (Equation 17) is determined from them:

$$s_s = \sqrt{\frac{\sum_{i=1}^{n} {v_i}^2}{n-6}}, \tag{17}$$

where n is the number of points in the local neighbourhood of the selected point.

Figure 14 shows a scan of a terrain with variable curvature, showing the difference between a plane and a second order surface representation of the terrain shape. The point cloud is colored based on standard deviation of deviation of each point from the surface in the direction of the third eigenvector. It can be seen that rounded shapes stand out when the terrain is fitted with the plane, but this representation is unsuitable for noise filtering. Because the second order surface adheres better to the terrain, noise can be removed even in areas with variable curvature.

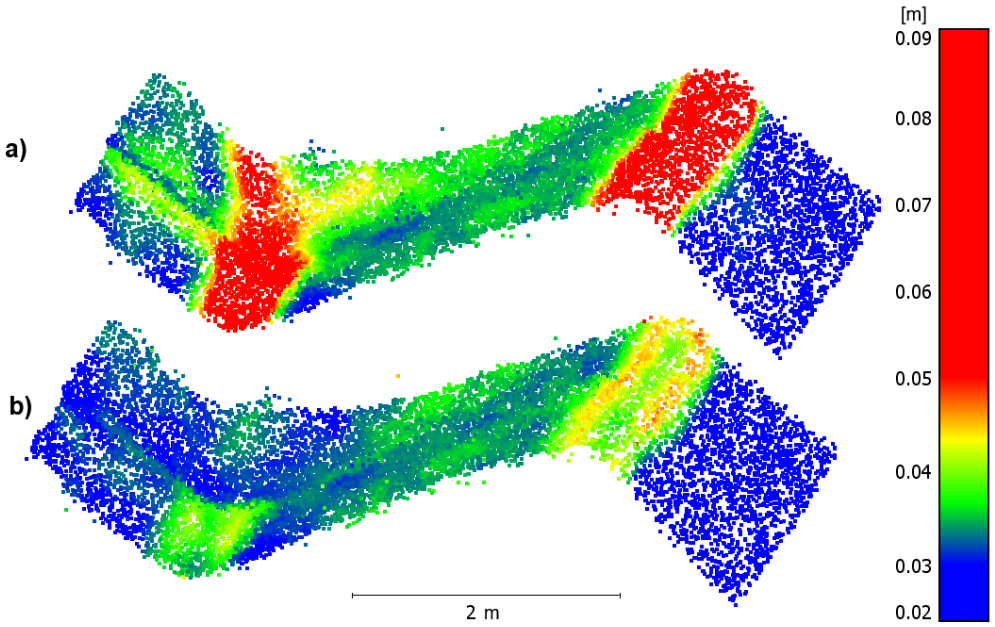

Fig. 14 – A point cloud of a terrain with variable curvature colored based on standard deviation of deviations from: a) a plane, b) a surface of the second order, both fitted to the points of the neighbourhood with radius of 0.5 m

Another limitation of the calculation of features from the local 3D neighbourhood of a point is the calculation method used, for example, in the CloudCompare software (paper [16] cited in their documentation [17]). The coordinates of the points falling within the 3D neighbourhood of the point of interest are reduced to the center of gravity of the neighbourhood points for computation purposes, and the eigenvalues and eigenvectors are then related to this center of gravity. This approach is suitable to describe the point cloud, or parts of it, as a whole. The reduction of a neighbourhood point to a center of gravity is done according to Equation 18:

$$X_{i,red} = X_i - \bar{X};\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \tag{18}$$

where X_i is a point of the neighbourhood, n is the number of points in the neighbourhood.

However, if the purpose of the feature is to determine the probability that a given point is noise, a different approach must be taken. If the point under study is part of the noise, it is often far from the center of gravity of the neighbourhood, which is likely to be located on the denser object we want to preserve. In order to characterize this point more accurately, it is possible to reduce the coordinates of the neighbourhood to the surveyed point to obtain information related directly to it. The reduction to the surveyed point is performed according to Equation 19:

$$X_{i,red} = X_i - X_P, \tag{19}$$

where X_P is the surveyed point.

Figure 15 shows the difference between the standard deviation (square root of the 3rd eigenvalue) to the center of gravity and to the surveyed point. The second approach is preferable for noise filtering.

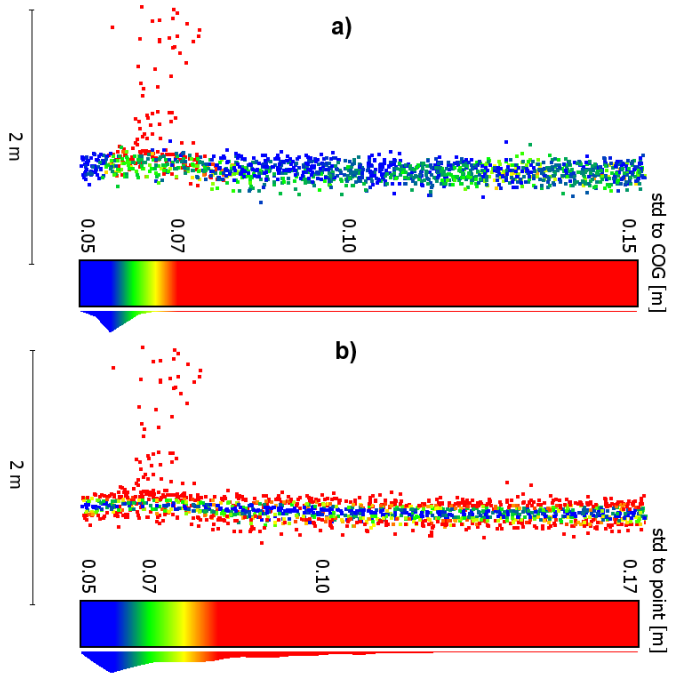

Fig. 15 – A point cloud of terrain with vegetation colored based on standard deviation of neighbourhood points in the direction of 3rd eigenvalue calculated from: a) the centre of gravity, b) the surveyed point

These two approaches can be further explored in the future and could increase the efficiency of automated filtering and classification of point clouds using neural networks and other types of machine learning.

CONCLUSION

This paper has provided a comprehensive review of the various features of the local 3D neighbourhood of a point in a point cloud that are used for data classification and analysis purposes. Furthermore, the suitability of using each feature to identify different types of objects including noise in the point cloud was evaluated.

Analysis of the dependencies between these features showed that some of them exhibit strong correlations with each other, suggesting the possibility of simplifying their use for data classification and filtering. For example, the strong correlation between anisotropy and sphericity indicates their interdependence. The use of only one of them can lead to an optimization of the selection of the most appropriate point cloud classification parameters.

Even though in the description of characteristics we reference several connections between the characteristics and the artificial or natural objects, we do not believe that these connections can be universally applied. The results of classification are always influenced by factors such as the nature of the data, the chosen scale or the level of noise.

New approaches proposed in this article could lead to more efficient classification and filtering of point clouds in the future. One of them is the use of a higher order surface for determining outliers in the neighbourhood. This approach could help remove noise in complex objects. Another new approach is to reduce the coordinates of the neighbouring points to the surveyed point rather than to the center of gravity of the neighbourhood. This makes it possible to obtain a characterization of the surveyed point with respect to its neighbourhood, not of the neighbourhood as a whole. These two new approaches should be the subject of further research.

ACKNOWLEDGMENTS

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS24/048/OHK1/1T/11 "Data filtering and classification using machine learning methods".

LIST OF SYMBOLS

3D – three-dimensional
X, Y, Z – point coordinates
RGB – red-green-blue color
PCA – principal component analysis
SVD – singular value decomposition
λ_i – i-th eigenvalue
COG – center of gravity
STD – standard deviation

REFERENCES

[1] PAVELKA JR., Karel; PACINA, Jan. Using of modern technologies for visualization of cultural heritage. Online. Stavební obzor - Civil Engineering Journal. 2023, vol. 32, No. 4, p. 549-563. ISSN 1805-2576. Available from: https://doi.org/10.14311/CEJ.2023.04.0041.
[2] MATOUŠKOVÁ, Eva; PAVELKA, Karel; SMOLÍK, Tobiáš and PAVELKA, Karel. Earthen Jewish Architecture of Southern Morocco: Documentation of Unfired Brick Synagogues and Mellahs in the Drâa-Tafilalet Region. Online. Applied Sciences. 2021, vol. 11, No. 4. ISSN 2076-3417. Available from: https://doi.org/10.3390/app11041712.
[3] KOVANIČ, Ľudovít; PEŤOVSKÝ, Patrik; TOPITZER, Branislav and BLIŠŤAN, Peter. Complex Methodology for Spatial Documentation of Geomorphological Changes and Geohazards in the Alpine Environment. Online. Land. 2024, vol. 13, No. 1. ISSN 2073-445X. Available from: https://doi.org/10.3390/land13010112.
[4] QI, Charles R.; SU, Hao; MO, Kaichun and GUIBAS, Leonidas J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Online. 2017. Available from: https://doi.org/10.48550/arXiv.1612.00593.
[5] ŠTRONER, Martin; URBAN, Rudolf and SUK, Tomáš. Filtering Green Vegetation Out from Colored Point Clouds of Rocky Terrains Based on Various Vegetation Indices: Comparison of Simple Statistical Methods, Support Vector Machine, and Neural Network. Online. Remote Sensing. 2023, vol. 15, No. 13. ISSN 2072-4292. Available from: https://doi.org/10.3390/rs15133254.
[6] ATIK, Muhammed Enesq and DURAN, Zaide. Classification of aerial photogrammetric point cloud using recurrent neural networks. Online. Fresenius environmental bulletin. 2021, vol. 30, No. 04A/2021, p. 4270-4275. ISSN 1018-4619. Available from: https://www.prt-parlar.de/download_list/?c=FEB_2021#.
[7] LIU, Kunbo; LIU, Shuai; TAN, Kai; YIN, Mingbo and TAO, Pengjie. ANN-Based Filtering of Drone LiDAR in Coastal Salt Marshes Using Spatial–Spectral Features. Online. Remote Sensing. 2024, vol. 16, No. 18. ISSN 2072-4292. Available from: https://doi.org/10.3390/rs16183373.
[8] CHEHATA, Nesrine; GUO, Li and MALLET, Clément. Airborne LIDAR feature selection for urban classification using random forests. Online. Laser scanning, IAPRS. 2009, Vol. XXXVIII. Available from: https://www.researchgate.net/publication/242731335_Airborne_LIDAR_feature_selection_for_urban_classification_using_random_forests.
[9] WEINMANN, M.; JUTZI, B. and MALLET, C. Feature relevance assessment for the semantic interpretation of 3D point cloud data. Online. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2013, vol. II-5/W2, p. 313-318. ISSN 2194-9050. Available from: https://doi.org/10.5194/isprsannals-II-5-W2-313-2013.
[10] WEINMANN, M.; JUTZI, B.; MALLET, C. and WEINMANN, M. GEOMETRIC FEATURES AND THEIR RELEVANCE FOR 3D POINT CLOUD CLASSIFICATION. Online. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2017, vol. IV-1/W1, p. 157-164. ISSN 2194-9050. Available from: https://doi.org/10.5194/isprs-annals-IV-1-W1-157-2017. [cit. 2024-11-23].

[11] KOVANIČ, Ľudovít; PEŤOVSKÝ, Patrik; TOPITZER, Branislav and BLIŠŤAN, Peter. Spatial Analysis of Point Clouds Obtained by SfM Photogrammetry and the TLS Method—Study in Quarry Environment. Online. Land. 2024, vol. 13, No. 5. ISSN 2073-445X. Available from: https://doi.org/10.3390/land13050614.
[12] RAPUCA, Almedina; MATOUŠKOVÁ, Eva. Testing of close-range photogrammetry and laser scanning for easy documentation of historical objects and buildings parts. Online. Stavební obzor - Civil Engineering Journal. 2023, vol. 32, No. 4, p. 504-518. ISSN 1805-2576. Available from: https://doi.org/10.14311/CEJ.2023.04.0038.
[13] WEINMANN, Martin; JUTZI, Boris; HINZ, Stefan and MALLET, Clément. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. Online. ISPRS Journal of Photogrammetry and Remote Sensing. 2015, vol. 105, p. 286-304. ISSN 09242716. Available from: https://doi.org/10.1016/j.isprsjprs.2015.01.016.
[14] JOLLIFFE, Ian T. and CADIMA, Jorge. Principal component analysis: a review and recent developments. Online. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2016, vol. 374, No. 2065. ISSN 1364-503X. Available from: https://doi.org/10.1098/rsta.2015.0202.
[15] BARNETT, T. P. and PREISENDORFER, R. Origins and Levels of Monthly and Seasonal Forecast Skill for United States Surface Air Temperatures Determined by Canonical Correlation Analysis. Monthly Weather Review. 1987, vol. 115, no. 9, p. 1825-1850. ISSN 0027-0644. Available from: https://doi.org/10.1175/1520-0493(1987)115<1825:OALOMA>2.0.CO;2.
[16] HACKEL, Timo; WEGNER, Jan D. and SCHINDLER, Konrad. Contour Detection in Unstructured 3D Point Clouds. Online. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016, p. 1610-1618. ISBN 978-1-4673-8851-1. Available from: https://doi.org/10.1109/CVPR.2016.178.
[17] Compute geometric features. Online. CLOUDCOMPARE. CloudCompare Wiki. 2024, 7. January 2024. Available from: https://www.cloudcompare.org/doc/wiki/index.php?title=Compute_geometric_features.