

Article

Progressive Dilution of Point Clouds Considering the Local Relief for Creation and Storage of Digital Twins of Cultural Heritage

Martin Štroner * , Tomáš Křemen and Rudolf Urban

Department of Special Geodesy, Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29 Prague, Czech Republic

* Correspondence: martin.stroner@fsv.cvut.cz

Abstract: Currently, the creation of digital copies (digital twins) of various objects by remote sensing methods producing point clouds is becoming commonplace. This might be particularly important for the digital preservation of historical objects. Such clouds are typically primarily acquired as unordered sets of points with regular dense spacing, making the clouds huge in size, which causes such clouds to be difficult to process, store and share. The clouds are, therefore, usually diluted before use, typically through uniform dilution with a set spacing; such dilution can, however, lead to the loss of detail in the resulting cloud (washed-out edges and fine features). In this paper, we present an easy-to-use and computationally inexpensive progressive dilution method preserving detail in highly rugged/curved areas while significantly reducing the number of points in flat areas. This is done on the basis of a newly proposed characteristic T , which is based on the local scattering of the cloud (i.e., on the ruggedness of the local relief). The performance of this algorithm is demonstrated on datasets depicting parts of historic buildings of different characters. The results are evaluated on the basis of (a) root mean square deviation (RMSD) between the original and diluted clouds, (b) of visual evaluation of the differences and (c) of reduction in the point cloud size, demonstrating an excellent performance of the algorithm with a minimum loss of detail while significantly reducing the point clouds (approx. by 47–66% compared to the corresponding uniform dilution for individual datasets).



Citation: Štroner, M.; Křemen, T.; Urban, R. Progressive Dilution of Point Clouds Considering the Local Relief for Creation and Storage of Digital Twins of Cultural Heritage. *Appl. Sci.* **2022**, *12*, 11540. <https://doi.org/10.3390/app122211540>

Academic Editor: Marek Fraštia

Received: 31 October 2022

Accepted: 11 November 2022

Published: 14 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: point cloud; subsampling; downsampling; dilution; cultural heritage; 3D scanning; SfM photogrammetry; point cloud size; detail

1. Introduction

Nowadays, remote sensing techniques, such as 3D scanning (e.g., [1,2]), structure from motion (SfM) photogrammetry [3,4], or UAV lidar [5] constitute common approaches for georeferencing and capture of historical objects. These methods produce huge amounts of data that can, among other uses, assist the documentation and preservation of cultural heritage objects for future generations [6,7] and/or the sharing of digital data on these objects among researchers or even with the public. Prior to processing, the dense point clouds collected by these techniques are typically diluted to a density that must correspond to the required detail. This leads to the need for compromise between the preservation of the most detailed parts of the object (such as reliefs or statues decorating a building) and the overall data size. If high detail is needed for the decorations, flat areas of such buildings (walls) would likely be oversampled, unnecessarily inflating the total data size. For this reason, a progressive dilution preserving a high point density on the parts needing high detail while reducing it for flat areas could be a promising approach for the storage of point clouds of such buildings.

Algorithms used for point cloud dilution can be generally classified into those working directly with point clouds, those using a triangular mesh for creating a surface, and other algorithms using, e.g., deep learning (neural networks).

Approaches directly diluting the point cloud are the simplest and most widely used. Algorithms performing uniform dilution are typical representatives of such approach, working, e.g., through random sampling (random selection of a user-defined number of points), octree sampling (the cloud point space is divided into voxels with user-defined size and one point—usually the closest one to the center of the voxel—is left in the cloud), or, the most widely used, space sampling method (points are excluded from the cloud on the basis of a user-defined minimal distance). These methods are very simple, fast, and implemented in the available software, e.g., in the CloudCompare software; thus, they are typically used as one of the first steps of point cloud processing ([1,8]).

The mesh-based approaches (e.g., [9–14]) build on the fact that real-world objects consist of surfaces and the point cloud only “samples” these surfaces. Combining the individual neighboring points into a network of triangles formed by neighboring points from the cloud then approximates the surface. Obviously, where the surface is more curved, more points are needed for its accurate description than where the surface is flat. The generation of such surfaces is, however, too computationally demanding and typically requires further manual editing. For these reasons, the use of these algorithms for historical buildings or objects is quite complicated, making their suitability for point cloud dilution in these applications rather questionable.

Other approaches use, for example, deep learning- or artificial intelligence-based algorithms [15,16]. Similarly to algorithms using mesh, however, their practical use is complicated by high computational demands; moreover, to the best of our knowledge, it has never been used for historical buildings. Simplification approaches, known especially from the field of computer graphics, are based on the preservation of only feature (key) points, such as edges, from the cloud, ensuring that the reduction causes a minimal change in the shape of the captured object (e.g., [17–21]). Specialized methods for the point cloud density reduction that have been originally designed for point clouds describing terrain- [22–24] could be also used for the description of cultural heritage objects.

A vast majority of the above-mentioned approaches for point cloud dilution has been developed for other purposes, such as small objects or terrain. On the other hand, the specific field of buildings with their associated characteristics (e.g., requirement for the target density of the processed point cloud in points per m^2) have been, so far, evaluated only in a few studies [25–27]. Approaches described in these papers are, moreover, relatively complicated and, in effect, computationally demanding.

For these reasons, we propose a simple and computationally undemanding method of progressive dilution that can be easily implemented in the freely available CloudCompare software v. 2.12.4 Kyiv (cloudcompare.org; 1 September 2022). This method is suitable for dense point clouds obtained through SfM photogrammetry or laser scanning capturing cultural heritage objects, such as buildings. The method is based on space sampling but the distance for exclusion of surrounding points differs, being calculated individually for each point based on the flatness of its surroundings expressed as a standard deviation of plane fit transformed into the variable T through a specifically designed transformation. The operator defines, based on the intended further use of the data, the required maximum and minimum dilution parameters (the distance between the neighboring points) for the least and most rugged surfaces.

The proposed method is tested on point clouds of three constructions (a bridge and two buildings) representing typical cultural heritage objects.

2. Materials and Methods

2.1. The Principle of the Method

The principle of the proposed method is quite simple. The point cloud representing the object's surface contains a finite number of discrete points. The surface of flatter parts can be captured using much fewer points than curved or rugged surfaces. If we can, therefore, determine the “flatness” of the surroundings of the point, we can determine how much we can dilute the cloud in its vicinity without compromising the quality of the

object description. The proposed method performs this in the unordered original point cloud using already known simple manipulations. Unlike the commonly used uniform dilution using only one parameter determining the distance between individual points in the diluted cloud (a), the proposed method adds an additional parameter (b), which defines the maximum distance between points, i.e., the distance that shall be used on the flattest surfaces. The method sequence is as follows (also shown on Figure 1):

1. Calculation of an evaluation variable E_3 .
2. Transformation of the E_3 variable into the T variable.
3. Correction of the range of the T variable by winsdorizing.
4. Dilution of the cloud based on a linearly changing distance between individual points based on the relative value of T .

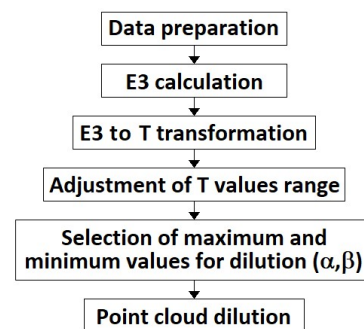


Figure 1. The flowchart of the proposed progressive dilution algorithm.

The E_3 variable is calculated separately for each point within the cloud. E_3 calculated for a point is defined as the variance of the distances of surrounding points (within a sphere with radius R) from the plane interspersed through them; it can be acquired, for example, as the third eigenvalue of the matrix according to Equation (1).

$$A = \begin{pmatrix} x_{1R} & y_{1R} & z_{1R} \\ \vdots & \vdots & \vdots \\ x_{nR} & y_{nR} & z_{nR} \end{pmatrix}, U\Sigma V = svd\left(\frac{A^T A}{n}\right), \text{diag}(\Sigma) = \begin{pmatrix} E_1^2 \\ E_2^2 \\ E_3^2 \end{pmatrix}, \quad (1)$$

$$x_{iR} = x_i - x_m, y_{iR} = y_i - y_m, z_{iR} = z_i - z_m,$$

$$x_m = \frac{\sum_{i=1}^n x_i}{n}, y_m = \frac{\sum_{i=1}^n y_i}{n}, z_m = \frac{\sum_{i=1}^n z_i}{n}.$$

The calculation of the eigenvalues of vectors/matrices is generally well-known; here, we present a procedure using a singular decomposition (function $svd()$) used in Equation (1). In that equation, A is a matrix of coordinates of points x_{iR}, y_{iR}, z_{iR} (i.e., the coordinates of each point within the spherical surroundings of the evaluated point with radius R , relative to the mean coordinates of all points within the sphere x_m, y_m, z_m).

Subsequently, the E_3 variable is transformed into the T variable. E_3 , as a characteristic of variance, greatly changes with even small changes in flatness; for this reason, we have proposed a transformation into the T variable (Equation (2)), which allows more linear changes in dilution.

$$T = \frac{1}{\sqrt{E_3}} \quad (2)$$

Moreover, in every cloud, there are extreme points excessively broadening the range of the T variable. A typical histogram, then, looks like that shown in Figure 2a—a vast majority of points fall within a relatively small range and only a small minority is outside this range. Values of T for individual points can be as high as $T = 35,000$ but there are so few such points that they are even not visible on the histogram. For this reason, winsdorizing is performed—the 99.9th percentile of data is arbitrarily found and this number is assigned to all points with T values above this percentile, see Figure 2b. This includes points on

extremely smooth and planar surfaces, for which T was undefined (E_3 for these points is 0 and T would, therefore, be $+\infty$; in CloudCompare, such points would be marked NaN). The same is done for the 0.1th percentile and points lower than 0.1. This is the reason why the $T = 3000$ category is seemingly so overrepresented in Figure 2b—it is not because the value of $T = 3000$ is valid for so many points but because all points with $T > 3000$ were aggregated into this single value.

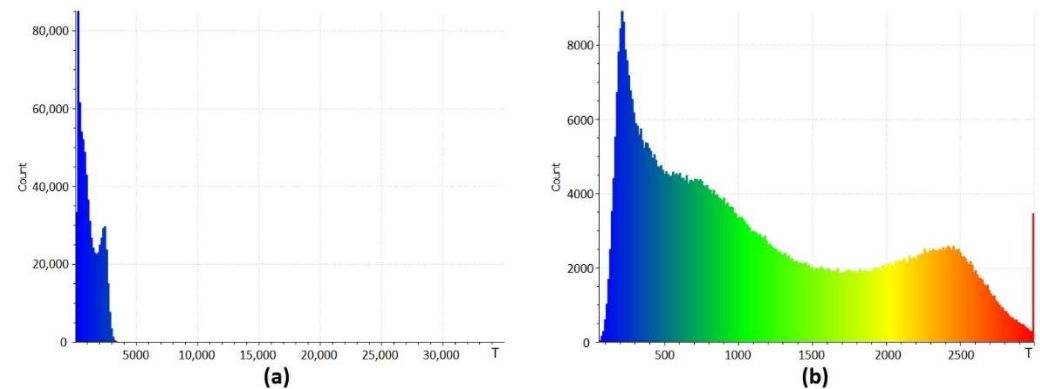


Figure 2. Histograms for the T variable (a) before and (b) after winsdorizing.

In the last step, the cloud is diluted based on the distances of the neighboring points. Unlike in uniform dilution, we do not use a universally valid distancing parameter; rather, two parameters are set, namely a as the minimum distance of neighboring points for the lowest T (i.e., for the most rugged surfaces) and b as the minimum distance of the neighboring points for the maximum T . Based on the T value for each point, the distance of the particular point from the neighboring ones is then determined by linear interpolation between these two parameters (a and b).

Of course, the data need to be standardly processed before the above-described method is used—in particular, it must be cleaned of points not characterizing the object (such as vegetation). Existing geometrical filters e.g., [28,29], can be used for this purpose; it must be, however, kept in mind that most such filters are originally designed for the extraction of terrain, and it might be beneficial to consider also other approaches based on the evaluation of data variance for the extraction of buildings [30,31].

2.2. Data Used for Method Testing

Three datasets (see Figure 3) were selected for testing, representing different historical objects with various characteristics. The data are neither perfect nor complete (best seen in Data 3) to evaluate if this represents an obstacle for the algorithm function.

Data 1 show a side of a historical bridge structure. The dataset was obtained photogrammetrically, the images were taken with a Canon EOS 400 D digital SLR camera and processed in Agisoft Metashape software v. 1.8.1. In total, the original cloud consists of 728,137 points, the resolution on average approx. 0.01 m. In terms of ruggedness, the central pillar is particularly pronounced (round at the bottom, square at the top). The water drains on the left and right sides and the ledges represent significant features that need to be preserved; the joints between the stones can be considered the smallest details. The data are colored based on real-world photo colors. This object is a typical candidate for digital twin creation—it represents a typical historical object with flat surfaces combined with spatial details that need to be preserved.

Data 2 and Data 3 have been selected as extremely complicated scenes to evaluate the performance of the proposed algorithm also for such difficult data. Data 2 describe a part of a chapel gradually falling into disrepair—its surface is formed by crumbling plastering revealing bricks underneath. The cloud was acquired using a terrestrial 3D scanner Leica P40. The total number of points is 407,532, the resolution is approx. 0.01 m. There are

practically no flat surfaces here, the surface is highly rugged from the perspective of both the original shape and the consequences of the decaying plaster. The data are colored based on the intensity of the reflected electromagnetic radiation (recorded by the 3D scanner).

Data 3 show the entrance gate of a historic town hall (after recent refurbishment). The cloud was acquired with the Leica P40 terrestrial 3D scanner, the total number of points is 1,470,249, the resolution is approximately 0.005 m. The data was included in testing because the shapes are sharp and undamaged, and the data combine areas of high detail and ruggedness (e.g., the coat of arms in the middle and areas such as the column footings, etc.) with relatively flat areas. There is, therefore, an extremely wide range of ruggedness. The data are colored based on the intensity of the reflected electromagnetic radiation (recorded by the 3D scanner).

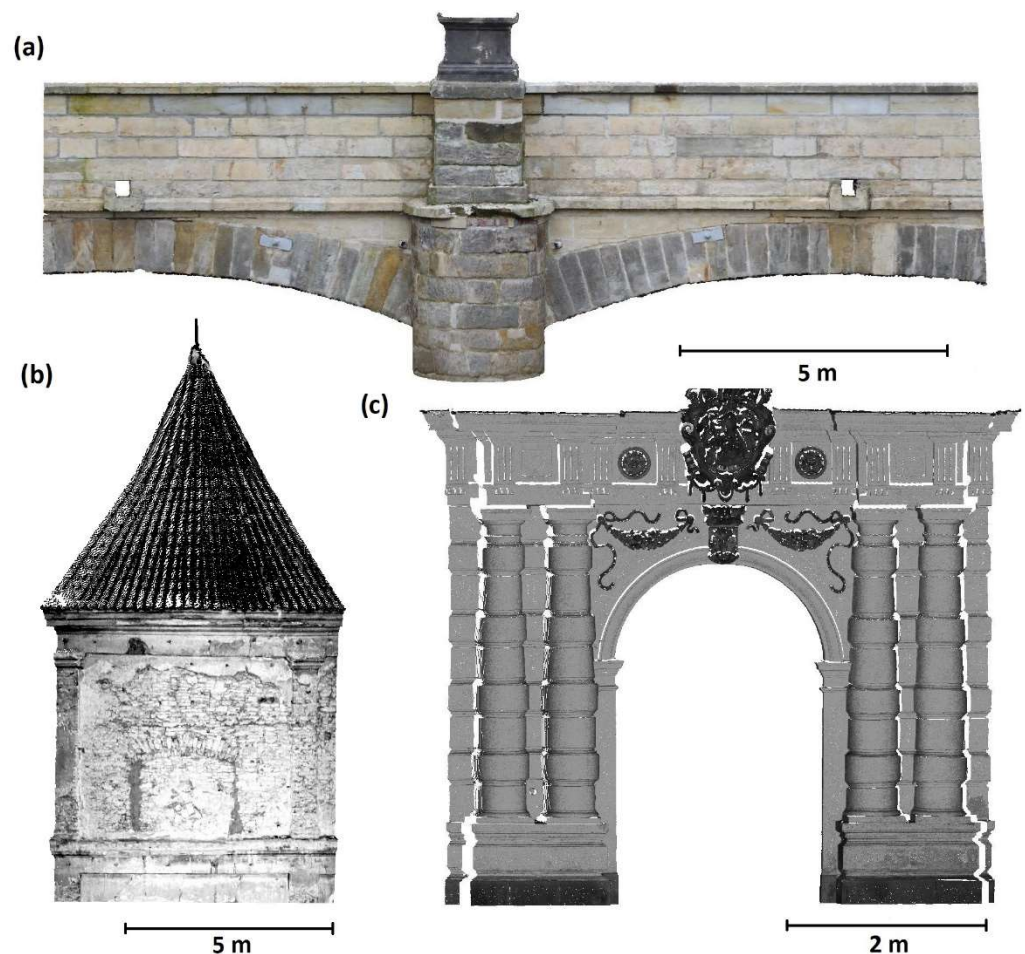


Figure 3. (a) Data 1—a part of a historic bridge; (b) Data 2—a part of a chapel falling into disrepair; (c) Data 3—a part of the entrance of the historic town hall.

2.3. Algorithm Performance Testing

The algorithm performance was tested by comparing the results of uniform dilution with those of the proposed progressive method. Evaluated parameters included the number of points in the respective cloud and the root mean square deviation (RMSD) between the local triangular mesh created from the diluted cloud and points from the original cloud (i.e., RMSD of the distances of the original points from the surface created from the diluted cloud). These two simple criteria indicate the quality of dilution—the best result of dilution contains the least number of points while being most similar to the original cloud (i.e., having the smallest RMSD between the original and diluted cloud). An additional parameter, $RMSD_E$, was also calculated after the exclusion of points that were not removed from the original cloud during dilution. The reason for this lies in the

fact that RMSD is calculated from all points in the cloud, (i.e., including such unremoved original points whose deviation from the original cloud must, by principle, be 0). This would falsely improve RMSD for clouds that were only slightly diluted and, therefore, contain a large number of such identical points. In other words, $RMSD_E$ is calculated only from the approximated surface, not from any original points.

2.3.1. Preparation of the Test Clouds

All testing was performed in CloudCompare v. 2.12.4 Kyiv (cloudcompare.org; 1 September 2022), a free 3D point cloud (and triangular mesh) processing software that was initially intended to compare two dense 3D point clouds or a point cloud and a triangular mesh. Over time, however, it has been expanded into a multipurpose point cloud processing program that includes a number of advanced algorithms (registration, resampling; handling of color, normals, scalar fields, statistical evaluations, etc.). It is widely used both for scientific purposes and for practical data processing. The calculation of E_3 values and its assignment into the scalar field was performed using the function Tools/Others/Compute geometric features, and the T variable was subsequently calculated using mathematical functions under Edit/Scalar fields/Arithmetics. The dilution itself (uniform as well as progressive dilution) was performed using the Edit/Subsample function, opting for the method Space. Subsequently, a constant parameter “min space between points” (set, e.g., to 0.01 m) was chosen in the case of uniform dilution; for progressive dilution, the option “use active SF” (SF = scalar field) is selected, and the a and b values are set for the maximum and minimum T .

2.3.2. Comparison of Prepared Clouds

The distances between the original point cloud and the triangular mesh created from the diluted cloud were calculated using the function Tools/Distances/Cloud to Cloud dist. Default settings were used, amended only in the Local modeling tab, where “Local model” 2D1/2 Triangulation was selected, and the Radius Sphere was set to $2 \times b$. The diluted point cloud was set as the reference cloud, the original as the compared cloud. RMSDs were subsequently determined using the function Tools/Statistics/Compute stat. params with Gaussian distribution.

Although RMSD is an exact parameter frequently used for cloud comparison, its information value is limited due to the immense number of points, which may lead to falsely good RMSD values despite inaccuracies in the filtered model; this is particularly true for points on the edges as the proportion of such is relatively low but their importance in preserving detail is high. For this reason, visual evaluation was employed as a useful additional criterion and will be presented in Results.

Reduction of the point cloud size is, of course, another important parameter. For this reason, we have also evaluated the reduction in the number of points in the diluted point cloud compared to the original one. In addition, we have also compared the reduction in the progressively diluted point cloud to the uniform dilution with the same a .

Table 1 presents the range of constants (values of R , a , b) used for testing. For each dataset, combinations of parameters were selected in view of the resolution of the original data and the level of detail that needed to be captured (for example, as Data 3 contain most details, a smaller initial a value was used). a and b values were proposed as (a) $b = 2a$; (b) $b = 3a$. Once these combinations were evaluated, some additional combinations were added arbitrarily based on these results—these are shown in Table 1 with designation (c). For Data 3, a lower initial $a = 0.005$ m was chosen due to the presence of high-detail reliefs in this dataset.

Table 1. The minimum distances of neighboring points used for dilution and testing.

Data	Uniform Dilution [m]	R [m]	a/b [m] for Progressive Dilutions
Data 1	0.01; 0.02; 0.03; 0.05	0.050	(a) 0.01/0.02; 0.02/0.04; 0.03/0.06; 0.05/0.10; (b) 0.01/0.03; 0.02/0.06; 0.03/0.09; 0.05/0.15; (c) 0.01/0.05; 0.02/0.10
Data 2	0.01; 0.02; 0.03; 0.05	0.100	(a) 0.01/0.02; 0.02/0.04; 0.03/0.06; 0.05/0.10; (b) 0.01/0.03; 0.02/0.06; 0.03/0.09; 0.05/0.15; (c) 0.01/0.10; 0.02/0.10;
Data 3	0.005; 0.01; 0.02	0.050	(a) 0.005/0.01; 0.01/0.02; 0.02/0.04; (b) 0.005/0.015; 0.01/0.03; 0.02/0.06; (c) 0.005/0.025; 0.005/0.050; 0.01/0.05; 0.02/0.10

3. Results

3.1. Testing Results—Data 1

At the first sight, Data 1 represents a simple object consisting of simple geometrical shapes. Figure 4 depicts the original data colored according to the T values. It is important to note that the T value distinguishes areas with different levels of ruggedness—this is the very principle of the method as without these distinguishing capabilities, the progressive dilution could not be satisfactory. Blue color indicates the most rugged/curved parts, red color the flattest ones. The figure clearly shows that the edges and joints are well distinguished from flat surfaces—including details such as the round blue dots in the lower part of the figure, indicating bolt heads. On a standard present-day laptop, the E_3 calculation for all points in this cloud took 1.2 s; the transformation of E_3 to T was practically immediate, and the dilution step took 0.4 s.

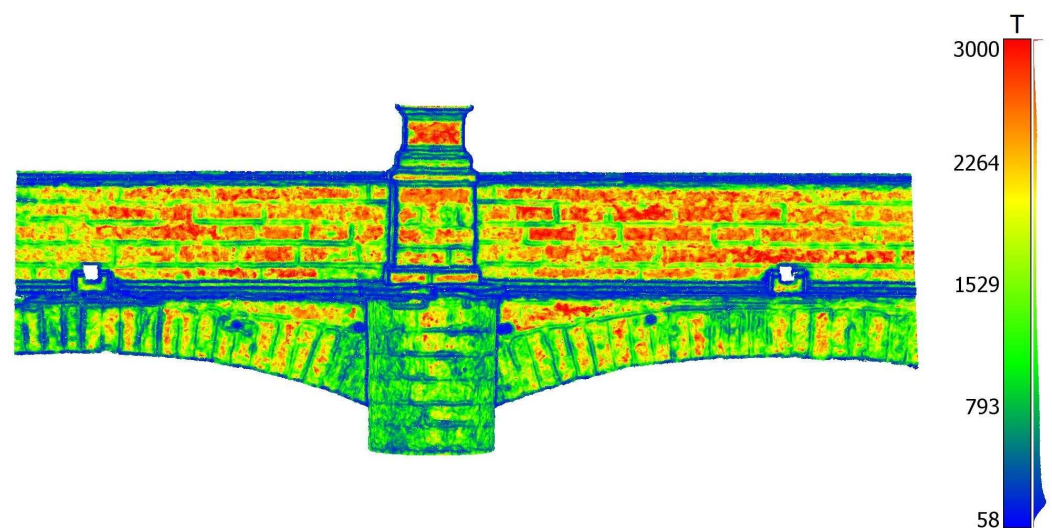


Figure 4. Data 1—The original point cloud colored according to the T values; red—flattest parts, blue—most curved/rugged parts.

Figure 5 indicates the results of processing with selected parameters and confirms that the progressive dilution works as intended. Figure 5a shows data after uniform dilution using $a = 0.02$ m, the remaining variants show the same data filtered using the same a but higher b , thus preserving details with the same accuracy as uniform dilution but lowering the numbers of points in flat areas.

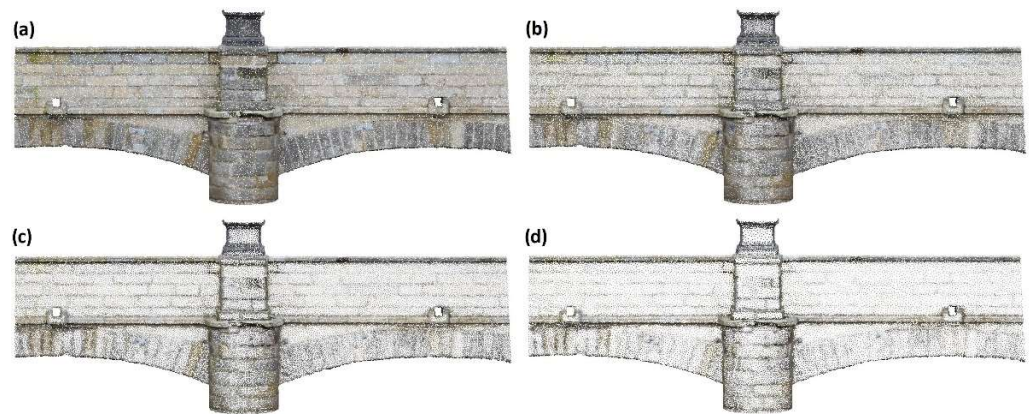


Figure 5. Examples of dilution: $a = 0.02$ m in all cases; (a) uniform dilution (i.e., $b = 0.02$ m); (b) $b = 0.04$ m; (c) $b = 0.06$ m; (d) $b = 0.10$ m.

Table 2 presents the results of the numerical testing of Data 1. As can be expected, increasing the a value leads to a decrease in the number of points in the diluted point cloud and the RMSD grows. Obviously, a is crucial for the capture of details as it sets the maximum possible density (maximum detail) of the resulting cloud. The quality of results should be, therefore, compared for the uniform and progressive dilutions with the same a . If, for example, the number of points after progressive dilution with settings of $a = 0.01$ m a $b = 0.02$ m is approximately 30% lower than that of uniform dilution while the RMSD remains practically unchanged (0.0005 m vs. 0.0006 m, respectively), it can be perceived as a successful application of the tested algorithm.

Table 2. Results of dilution of the Data 1 cloud—numbers of points and RMSDs ($R = 0.05$ m).

a [m]	b [m]	Points in the Cloud	% Original ¹	RMSD [m]	RMSD _E [m]	% Uniform ²
original	-	728,137	100	-	-	-
0.01	uniform	352,660	48.4	0.0005	0.0007	100
0.01	0.02	249,549	34.3	0.0006	0.0007	70.8
0.01	0.03	187,024	25.7	0.0006	0.0007	53.0
0.01	0.05	120,506	16.5	0.0007	0.0008	34.2
0.02	uniform	121,333	16.7	0.0010	0.0011	100.0
0.02	0.04	81,622	11.2	0.0011	0.0012	67.3
0.02	0.06	59,777	8.2	0.0012	0.0013	49.3
0.02	0.10	36,728	5.0	0.0016	0.0016	30.3
0.03	uniform	65,148	8.9	0.0015	0.0016	100.0
0.03	0.06	40,194	5.5	0.0017	0.0017	61.7
0.03	0.09	29,088	4.0	0.0020	0.0020	44.6
0.05	uniform	26,194	3.6	0.0029	0.0030	100.0
0.05	0.10	15,752	2.2	0.0033	0.0033	60.1
0.05	0.15	11,100	1.5	0.0042	0.0042	42.4

¹ % original—the percentage of points relative to the original cloud. ² % uniform—the percentage of points relative to the uniformly diluted cloud with respective a .

The detail presented in Figure 6e allows a better visual evaluation of the performance of the progressive dilution algorithm. Clearly, the density of points remains much higher on the edges and in the joints between stones than on the flat sides of the stones. The edges of details, such as the bolt head in the lower left corner are well preserved, but their flat parts are successfully diluted.

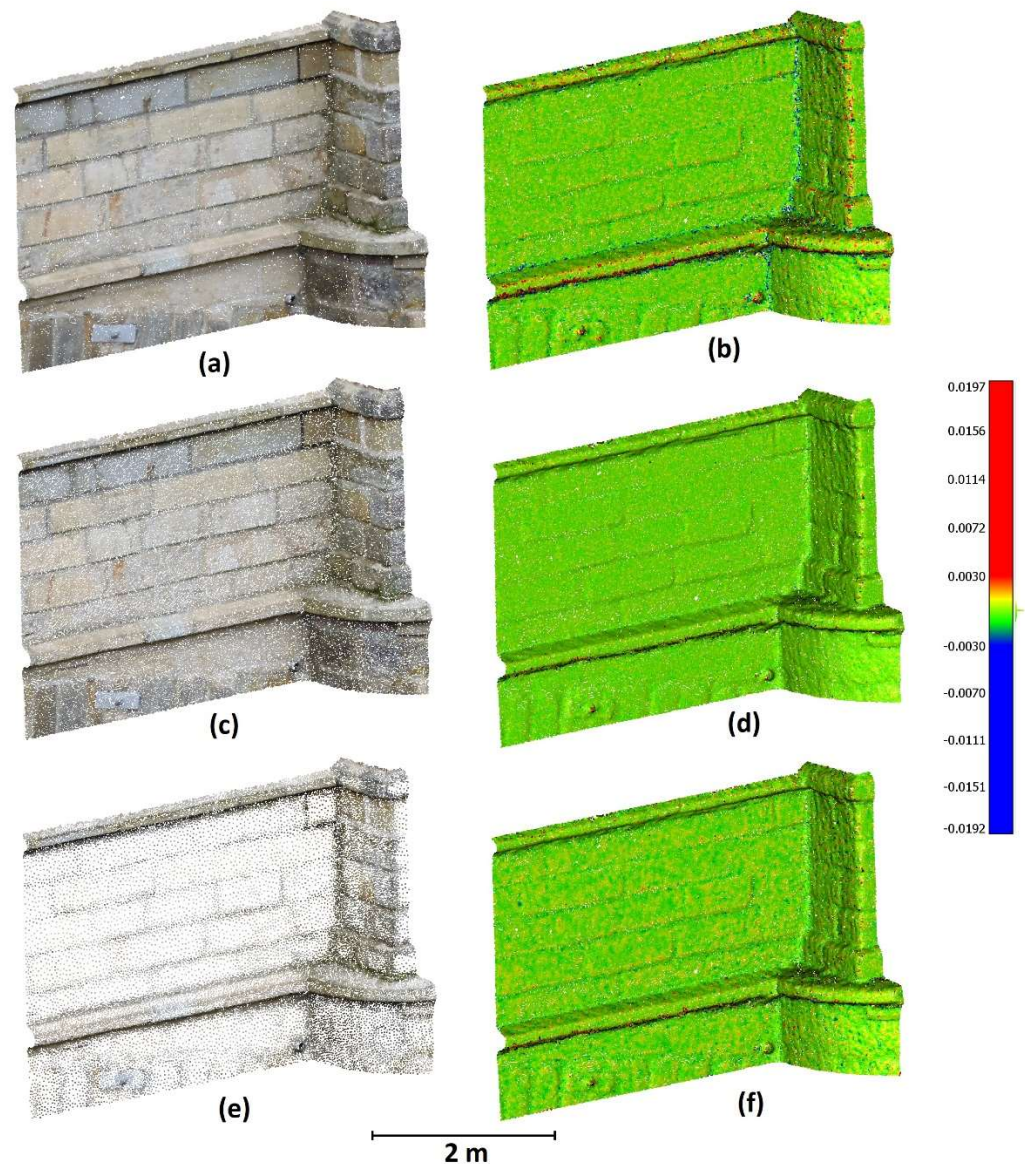


Figure 6. Details of the Data 1 point cloud: (a) original cloud, (b) differences between the cloud acquired by uniform dilution with $a = 0.02$ m and the original point cloud, (c) point cloud after uniform dilution with $a = 0.01$ m, (d) differences between the cloud acquired by uniform dilution with $a = 0.01$ m and the original point cloud, (e) point cloud after $a = 0.01$ m $b = 0.05$ m progressive dilution (approximately the same number of points in the cloud as for uniform dilution 0.02 m), (f) differences between the cloud acquired by progressive dilution with $a = 0.01$ m $b = 0.05$ m and the original point cloud. Note the higher representation of deviations on the edges of (b) compared to (f).

Figure 6 shows a detail of the point clouds diluted using various settings and their differences from the original cloud. Note in particular the differences between Figure 6b,f—two point clouds containing approximately the same numbers of points after dilution (Table 2). Figure 6b compares the original cloud with the result of uniform dilution with $a = 0.020$ m while (f) with the result of progressive dilution using $a = 0.010$ m and $b = 0.050$ m. The progressive cloud has a better RMSD (Table 2), which corresponds to the visual comparison where greater differences on the edges from the original cloud can be observed for the uniform dilution with higher alpha (red and blue points on the edges). This illustrates the main advantage of the progressive dilution—by using higher density only locally, it preserves greater detail only where needed, thus reaching better results with the same final cloud size.

Comparing Figure 6d,f we can see the comparison of two clouds with the same maximum detail a (0.010 m). The progressively diluted cloud contains approx. a third of points compared to the uniform dilution, which represents a significant compression of the data. Although the number of points with a greater difference from the original cloud (red) is somewhat higher in the progressively diluted cloud, we can see that such erroneous points are dissipated, representing largely chinks and slight unevenness in the relatively flat areas of stone faces while the edges and details are preserved very well.

3.2. Testing Results—Data 2

Data 2 represent an object that is practically free of flat surfaces. Figure 7 shows the point cloud colored according to the T values (flat surfaces in red, rugged surfaces in blue). Rather than architectural decorations, the high detail originates from the defects, chinks, and cracks in this object. However, even these are important for an accurate recording of the object. A closer look reveals also ledges and other projections. On a standard laptop, the E_3 calculation took less than 1 s, its transformation to T was practically immediate and the dilution took 0.2 s.

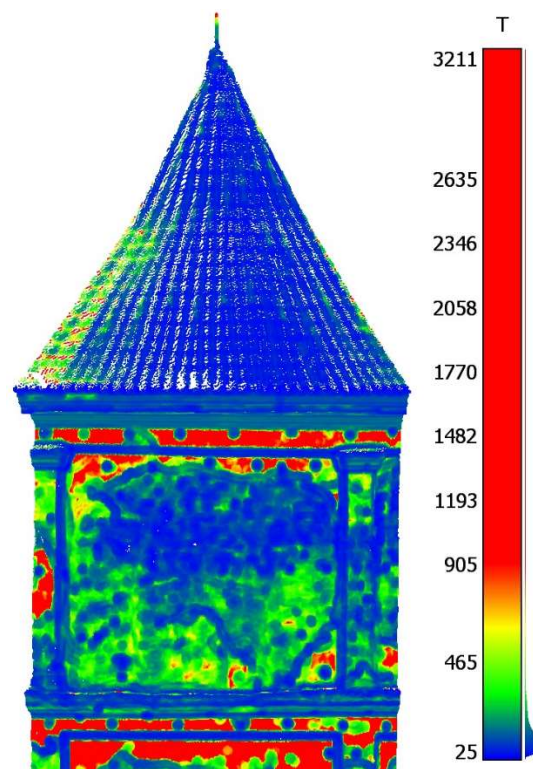


Figure 7. Data 2—original point cloud colored according to the T values; red—flattest parts, blue—most curved/rugged parts.

Figure 8 shows examples of the processed cloud for the walled part of the chapel with $a = 0.02$ m and different values of b . Table 3 shows the numerical values indicating the reduction in the number of points; using the same example as in Data 1, i.e., $a = 0.01$ m and $b = 0.02$ m, 0.03 m or 0.05 m, the number of points in the cloud decreased to 90%, 82% and 70%, respectively.

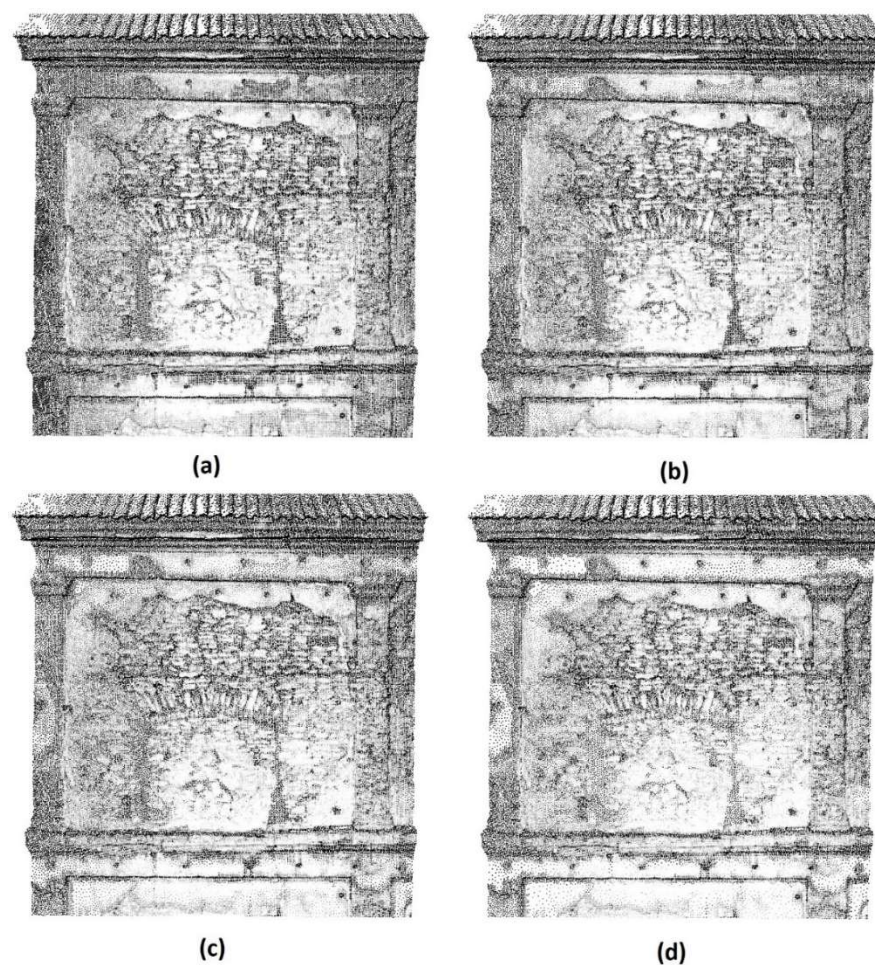


Figure 8. Data 2—examples of dilution: $a = 0.02$ m in all cases; (a) uniform dilution (i.e., $a = 0.02$ m); (b) $b = 0.04$ m; (c) $b = 0.06$ m; (d) $b = 0.10$ m.

Table 3. Results of dilution of the Data 2 cloud—numbers of points and RMSDs ($R = 0.10$ m).

a [m]	b [m]	Points in the Cloud	% Original ¹	RMSD [m]	RMSD _E [m]	% Uniform ²
original	-	407,532	100	-	-	-
0.01	uniform	404,419	99.2	0.0001	0.0015	100.0
0.01	0.02	366,731	90.0	0.0002	0.0006	90.7
0.01	0.03	332,062	81.5	0.0003	0.0008	82.1
0.01	0.05	283,446	69.6	0.0005	0.0010	70.1
0.01	0.10	213,828	52.5	0.0009	0.0013	52.9
0.02	uniform	120,139	29.5	0.0020	0.0023	100.0
0.02	0.04	112,243	27.5	0.0020	0.0024	93.4
0.02	0.06	105,751	25.9	0.0021	0.0025	88.0
0.02	0.1	95,120	23.3	0.0022	0.0025	79.2
0.03	uniform	63,517	15.6	0.0033	0.0036	100.0
0.03	0.06	58,674	14.4	0.0034	0.0037	92.4
0.03	0.09	54,882	13.5	0.0035	0.0038	86.4
0.05	uniform	26,158	6.4	0.0058	0.0060	100.0
0.05	0.10	23,756	5.8	0.0061	0.0063	90.8
0.05	0.15	22,082	5.4	0.0062	0.0064	84.4

¹ % original—the percentage of points relative to the original cloud. ² % uniform—the percentage of points relative to the uniformly diluted cloud, with respective a value.

The extremely low value of RMSD in Table 3 for $a = 0.01$ (0.1 mm for uniform dilution) is caused by the fact that in this dilution, practically no points were removed from the origi-

nal cloud (only 0.7% of points were removed) due to the original resolution being approx. 0.01 m. For this reason, evaluation using RMSDE gives a complementary perspective on the success of the dilution.

Figure 9 then shows a detail of Data 2, where (a) is a detail of the original cloud and (e) of the progressively diluted cloud. Again, we can see a lower density of points in the flat surfaces while density is preserved in the rugged parts such as the visible stones, bricks, or edges.

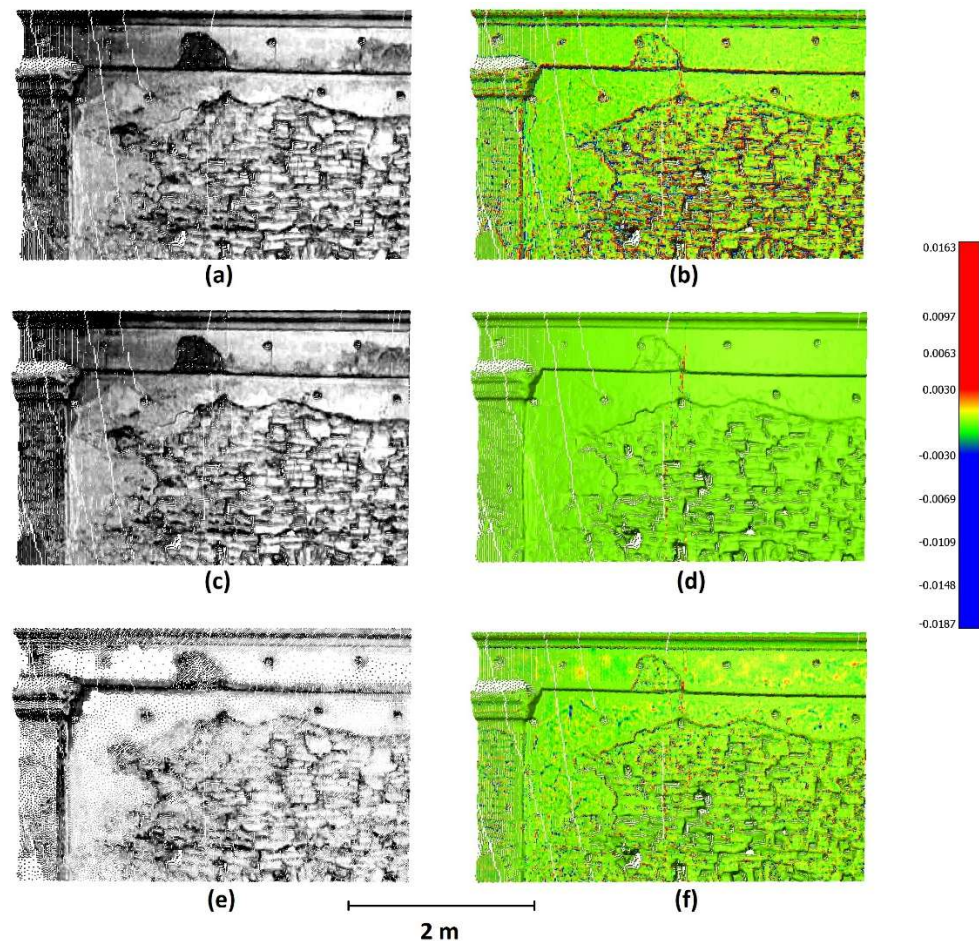


Figure 9. Details of the Data 2 point cloud: (a) original cloud, (b) differences between the cloud acquired by uniform dilution with $a = 0.02$ m and the original point cloud, (c) point cloud after uniform dilution with $a = 0.01$ m, (d) differences between the cloud acquired by uniform dilution with $a = 0.01$ m and the original point cloud, (e) point cloud after $a = 0.01$ m $b = 0.10$ m progressive dilution (approximately the same number of points in the cloud as for uniform dilution 0.02 m), (f) differences between the cloud acquired by progressive dilution with $a = 0.01$ m $b = 0.10$ m and the original point cloud.

Similarly to the details presented for Data 1, we can again see that increasing a leads to a major loss in detail, and although uniform dilution with the same a provides visually slightly better results than progressive dilution with 10 times higher b , it reduces the number of points by about half while preserving sufficient detail even in data of such complexity.

3.3. Testing Results—Data 3

Compared to the previous two datasets, the character of Data 3 differs in the sense of having a lot of details with sharp edges, which calls for a higher quality of the resulting model (diluted point cloud). The original resolution of the point cloud is also higher than in the previous cloud (better than 0.005 m). The data includes highly detailed reliefs, such

as the coat of arms in the middle and decorations of the columns, as well as practically flat surfaces on the feet of the columns or directly on the building walls. The areas covering the cylindrical parts of columns are slightly curved. Figure 10 shows the original point cloud colored according to the T variable. The E_3 calculation on a standard laptop took 2.2 s, the transformation of E_3 to T is practically immediate, and the dilution took 0.9 s.

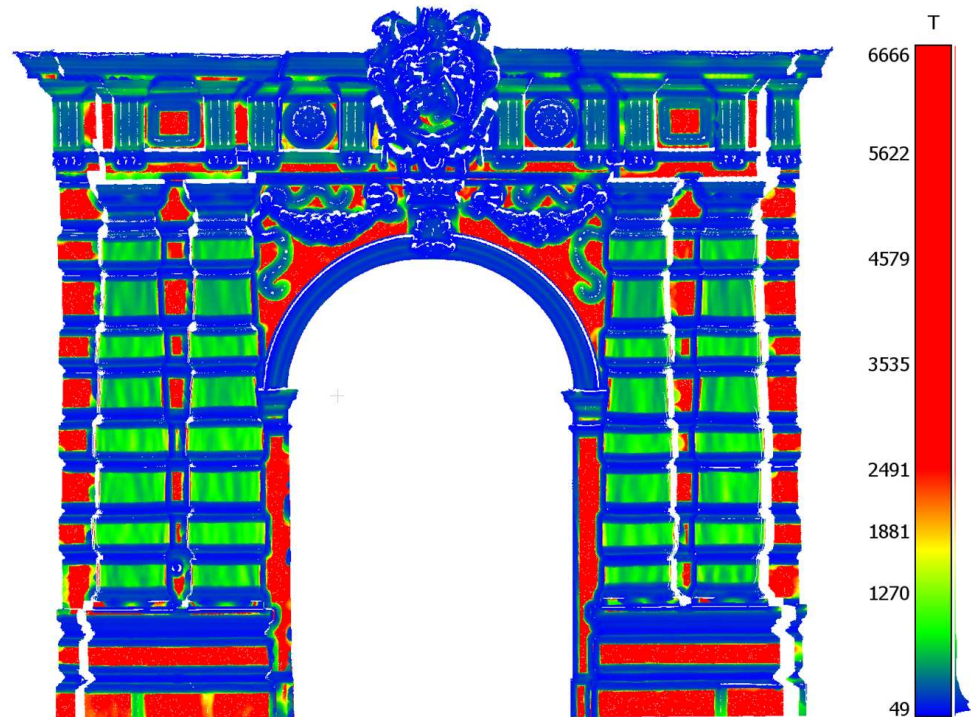


Figure 10. Data 3—original point cloud colored according to the T values; red—flattest parts, blue—most curved/rugged parts.

The results of dilution in Data 3 are similar to the other datasets (see Figure 11 and Table 4). Figure 12 shows a detail of the Data 3 cloud. The detail contains areas of various flatness/ruggedness/curvature (flat wall, cylindrical parts of the columns with various radiuses, and very dense areas representing fine details in the stucco reliefs).

Figure 12e clearly shows the level to which the flat surfaces can be diluted without loss of detail when using progressive dilution on well-preserved objects with plaster. Again, the loss of detail in complicated areas is most pronounced when a is increased (Figure 12b). Comparing the uniform dilution and progressive dilution with the same a , we can see only a very slight decrease in detail accuracy while reducing the point cloud size by more than 50%.

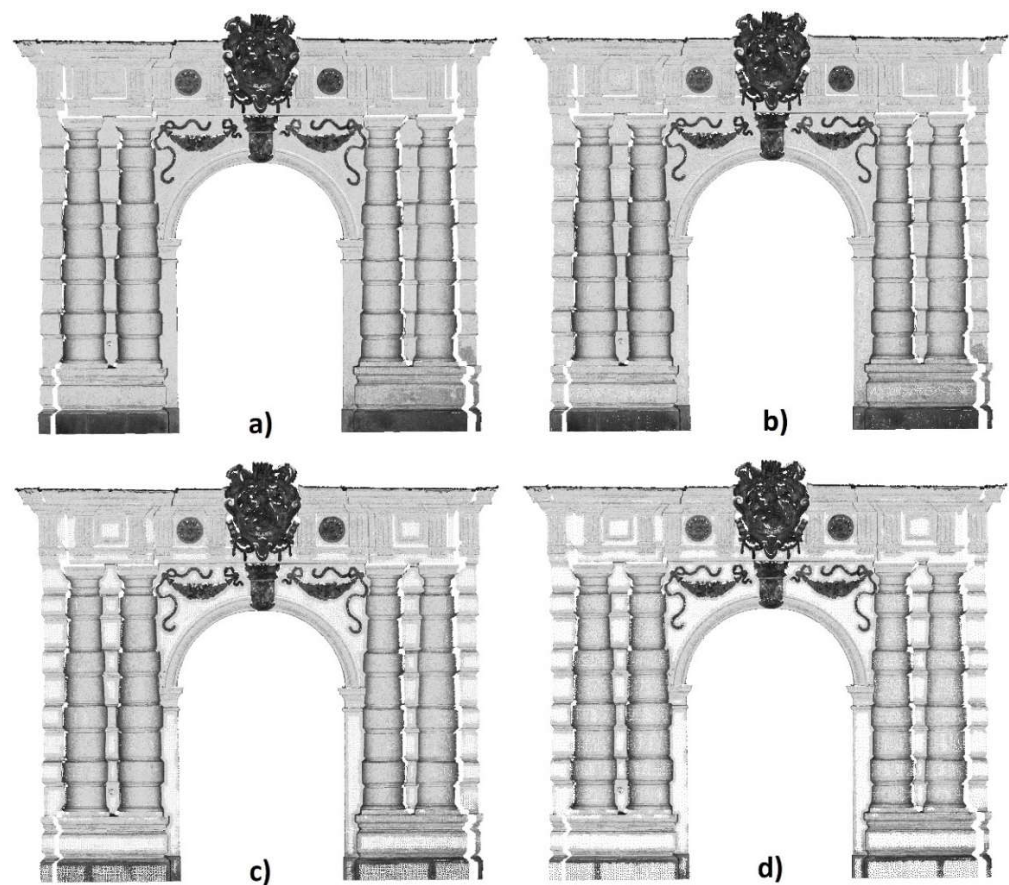


Figure 11. Data 3—examples of dilution: $a = 0.005$ m in all cases; (a) uniform dilution; (b) $b = 0.010$ m; (c) $b = 0.025$ m; (d) $b = 0.050$ m.

Table 4. Results of dilution of the Data 3 cloud—numbers of points and RMSDs ($R = 0.05$ m).

a [m]	b [m]	Points in the Cloud	% Original ¹	RMSD [m]	RMSD _E [m]	% Uniform ²
original	-	1,470,249	100	-	-	-
0.005	uniform	651,345	44.3	0.0004	0.0005	100.0
0.005	0.010	506,966	34.5	0.0004	0.0005	77.8
0.005	0.015	436,672	29.7	0.0004	0.0005	67.0
0.005	0.025	369,672	25.1	0.0005	0.0006	56.8
0.005	0.050	291,500	19.8	0.0006	0.0007	44.8
0.01	uniform	181,235	12.3	0.0009	0.0009	100.0
0.01	0.02	156,035	10.6	0.0009	0.0010	86.1
0.01	0.03	140,215	9.5	0.0010	0.0010	77.4
0.01	0.05	119,166	8.1	0.0011	0.0011	65.8
0.02	uniform	49,600	3.4	0.0021	0.0021	100.0
0.02	0.04	42,539	2.9	0.0022	0.0022	85.8
0.02	0.06	38,153	2.6	0.0024	0.0024	76.9
0.02	0.10	31,709	2.2	0.0046	0.0047	63.9

¹ % original—the percentage of points relative to the original cloud. ² % uniform—the percentage of points relative to the uniformly diluted cloud, with respective a value.

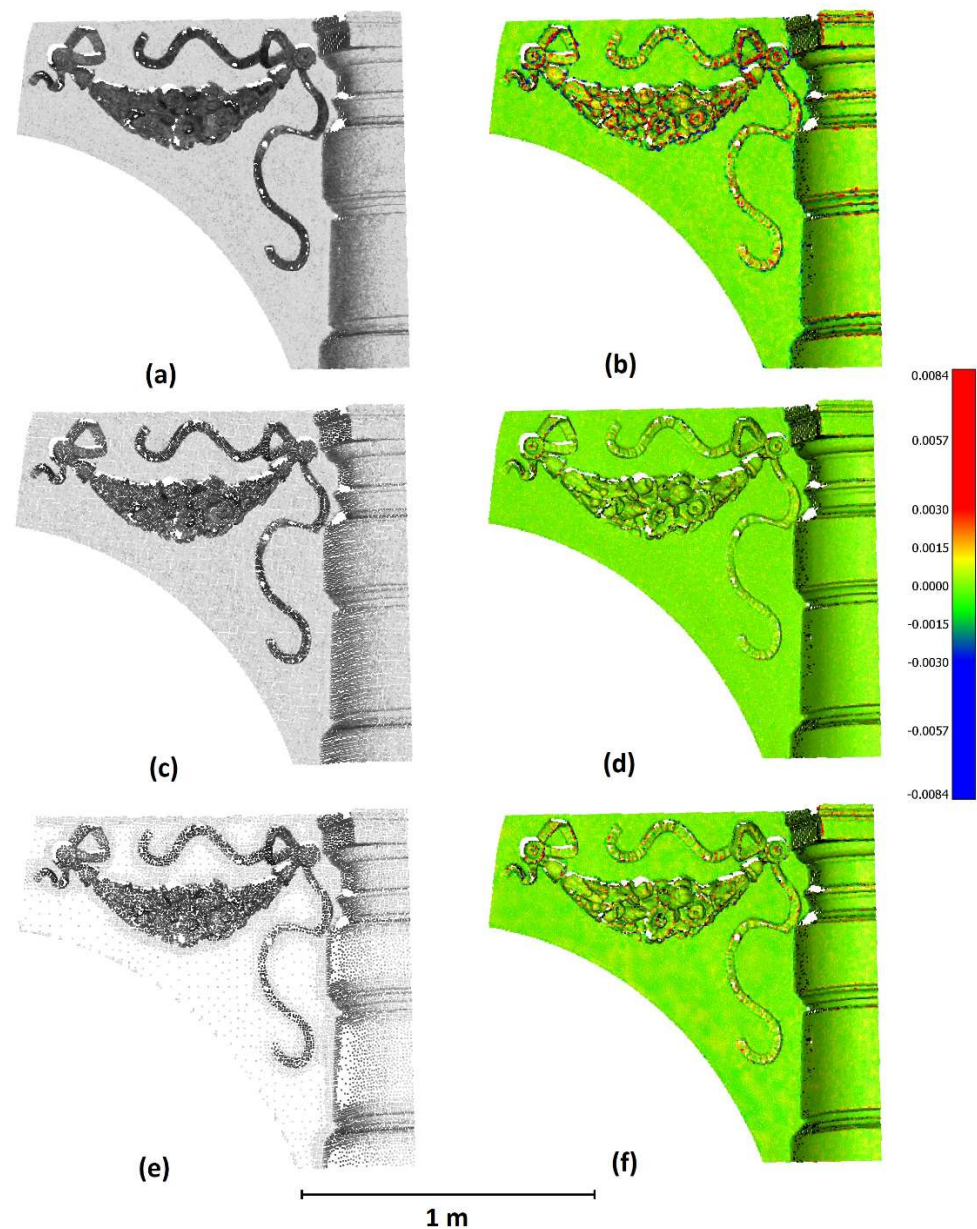


Figure 12. Details of the Data 3 point cloud: (a) original cloud, (b) differences between the cloud acquired by uniform dilution with $a = 0.01$ m and the original point cloud, (c) point cloud after uniform dilution with $a = 0.005$, (d) differences between the cloud acquired by uniform dilution with $a = 0.005$ m and the original point cloud, (e) point cloud after $a = 0.005$ m $b = 0.05$ m progressive dilution, (f) differences between the cloud acquired by progressive dilution with $a = 0.005$ m $b = 0.05$ m and the original point cloud.

4. Discussion

The main purpose of this algorithm for progressive dilution is to reduce the point cloud size while minimizing the loss of detail. The results prove that a is the key parameter for detail preservation—this is obvious both from the visual and numerical comparisons.

Three datasets were selected for the evaluation of the performance of the proposed algorithm. Besides the Data 1 dataset representing a typical medieval construction built from stones, we also used two highly complex structures. In one of those, the complexity and high level of detail are caused by the fact that the building is round and falling into disrepair, with the plaster falling from the walls forming irregular details (Data 2). In the other dataset, we can see a combination of flat surfaces (caused by the fact that the

surface has been relatively recently refurbished) and highly detailed reliefs and decorations (Data 3).

The reduction of the point cloud was, logically, the highest (66%) in the case of the medieval stone construction (Data 1), which contained only a limited number of details. On the other hand, a reduction of approximately only 48% was achieved for the best-performing setting in the case of Data 2 and 55% in the case of Data 3. This is not surprising, because the latter two datasets contain a high level of detail. Still, we believe that the achieved reduction at the expense of only a minor loss of detail can be considered a great success of the proposed algorithm. This is especially true if comparing the loss of detail experienced if attempting to reach the same reduction in the point cloud size using the uniform dilution, compare, e.g., Figure 6b,f. In the case of plastered buildings without fine embellishments, we could assume an even greater reduction in the point cloud size.

The numerical comparison of the dataset performance using RMSD also shows that progressive dilution leads only to a slight reduction in accuracy. We must, however, keep in mind that RMSD may not be the perfect parameter in this case—the number of points defining the edges and details is on most buildings is relatively low compared to that of those characterizing flat surfaces. In such cases, this numerical parameter is greatly influenced by these points on flat surfaces with practically 100% accuracy and although the final RMSD can be excellent, the fine edges and decorations may become “washed-out”. For this reason, we propose to always use, in addition to RMSD, also visual evaluation. For example, the RMSDs describing point clouds used for details in Figure 12b,f are very close and both excellent, below 1 mm (0.0009 and 0.0006 m, respectively), but the visual evaluation detects a great difference in detail preservation between these two clouds.

It must be also noted that if dilution does not lead to a significant reduction in the number of points in the cloud (see Table 3 and point cloud reduction and RMSDs for $a = 0.01$ m), RMSD is extremely good just because of the very fact that the point clouds are almost identical. For this reason, we introduced the RMSD_E parameter, which removes points that are identical in both point clouds (have not been removed) from the calculation. This plays no major role where the dilution was successful, as the number of these original points is relatively low. In such cases, the differences between the positions of the points from the original cloud and the mesh created from the diluted point cloud prevail and RMSD can provide a relatively good evaluation of the dilution accuracy.

As well as its good performance in reducing the point clouds and preserving details, the proposed progressive dilution algorithm comes with the additional benefits of simplicity (and, thus, rapidity of calculation) and availability (i.e., it can be easily implemented using the free CloudCompare software. For this reason, we believe that it has great potential to be used for the creation and sharing of digital twins of historical buildings. As progressive dilution would likely perform even better on modern constructions without embellishments, the algorithm is likely to be suitable not only for historical buildings but also for any other constructions and even contemporary civil engineering or for the creation of digital twins of entire cityscapes.

It is, however, necessary to draw attention to the basic prerequisite for the proper functioning of the algorithm—well-prepared data. The algorithm is based on the detection of data scattering from the fitted plane. For this reason, input data must be well cleaned of outliers arising due to errors in data registration, errors caused by a suboptimal combination of several scans (i.e., when the overlay of two scans creates two parallel surfaces) or suboptimal processing of the photogrammetric method producing noisy clouds. Such noise/parallel surfaces would inherently lead to an incorrect assessment of flatness (i.e., such areas would be considered areas of high detail) and progressive dilution would fail. Therefore, operator supervision and evaluation of the input cloud are required. This is, however, a common problem for all dilution algorithms—once any algorithm is fed with poor data, the result is bound to be poor as well.

5. Conclusions

Modern surveying methods such as laser scanning and photogrammetry make it possible to create nearly flawless and highly detailed digital 3D representations of objects of all forms and sizes, which allows the preservation of the objects of cultural or technical heritage in the digital form for future generations. The level of preserved detail may, however, become an obstacle for such preservation as, naturally, it increases the data size.

For this reason, we proposed a novel method for the progressive dilution of such point clouds, primarily intended for creating digital twins of historical buildings (but other applications are also possible). This algorithm is based on the determination of the flatness in the area around each point in the cloud (this is determined by fitting the plane through the spherical surroundings of the point and calculating the deviation of these points from the plane; where the surfaces are planar, this deviation would be minimum, while on curved/rugged surfaces, this deviation is large). Depending on this flatness and user-defined parameters of minimum and maximum dilution, the algorithm then dilutes flatter areas more than highly detailed ones, which brings a notable reduction in the data size compared to the uniform dilution, providing comparable detail. The advantages of this algorithm include a significant point cloud reduction with minimum loss of detail (demonstrated both numerically using RMSD and visually), easy implementation in free software CloudCompare, and rapidity of processing. The only limitation we are aware of is the need for good preparation of the input data and cleaning it from noise and/or data artifacts.

Author Contributions: Conceptualization, M.Š.; methodology, M.Š.; software, M.Š.; validation, R.U. and T.K.; formal analysis, M.Š. and R.U.; investigation, T.K.; writing—original draft preparation, M.Š.; writing—review and editing, R.U. and T.K.; visualization, M.Š.; funding acquisition, R.U. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Grant Agency of CTU in Prague—grant number SGS22/046/OHK1/1T/11, “Optimization of acquisition and processing of 3D data for purpose of engineering surveying, geodesy in underground spaces and 3D scanning”, and by the Technology Agency of the Czech Republic—grant number CK03000168, “Intelligent methods of digital data acquisition and analysis for bridge inspections”.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pérez-Álvarez, R.; de Luis-Ruiz, J.M.; Pereda-García, R.; Fernández-Maroto, G.; Malagón-Picón, B. 3D Documentation with TLS of Caliphal Gate (Ceuta, Spain). *Appl. Sci.* **2020**, *10*, 5377. [\[CrossRef\]](#)
2. Matoušková, E.; Pavelka, K.; Smolík, T.; Pavelka, K. Earthen Jewish Architecture of Southern Morocco: Documentation of Unfired Brick Synagogues and Mellahs in the Drâa-Tafilalet Region. *Appl. Sci.* **2021**, *11*, 1712. [\[CrossRef\]](#)
3. Ozimek, A.; Ozimek, P.; Skabek, K.; Łabędź, P. Digital Modelling and Accuracy Verification of a Complex Architectural Object Based on Photogrammetric Reconstruction. *Buildings* **2021**, *11*, 206. [\[CrossRef\]](#)
4. Marčiš, M.; Fraštia, M. Photogrammetric Measurement of a Wooden Truss. *Slovak J. Civ. Eng.* **2018**, *26*, 1–10. [\[CrossRef\]](#)
5. Roiha, J.; Heinari, E.; Holopainen, M. The Hidden Cairns—A Case Study of Drone-Based ALS as an Archaeological Site Survey Method. *Remote Sens.* **2021**, *13*, 2010. [\[CrossRef\]](#)
6. Niccolucci, F.; Felicetti, A.; Hermon, S. Populating the Data Space for Cultural Heritage with Heritage Digital Twins. *Data* **2022**, *7*, 105. [\[CrossRef\]](#)
7. Callieri, M.; Dellepiane, M.; Pavoni, G.; Pingi, P.; Potenziani, M.; Scopigno, R. Alchemy in 3D: A digitization for a journey through matter. In Proceedings of the International Congress on Digital Heritage, Granada, Spain, 28 August–2 October 2015; pp. 223–231.
8. Błaszczak-Bąk, W.; Janicka, J.; Suchocki, C.; Masiero, A.; Sobieraj-Żłobińska, A. Down-Sampling of Large LiDAR Dataset in the Context of Off-Road Objects Extraction. *Geosciences* **2020**, *10*, 219. [\[CrossRef\]](#)

9. Boltcheva, D.; Lévy, B. Surface Reconstruction by Computing Restricted Voronoi Cells in Parallel. *Comput.-Aided Des.* **2017**, *90*, 123–134. [[CrossRef](#)]
10. Hanocka, R.; Metzger, G.; Giryes, R.; Cohen-Or, D. Point2Mesh. *ACM Trans. Graph.* **2020**, *39*, 126.1–126.12. [[CrossRef](#)]
11. Williams, F.; Schneider, T.; Silva, C.; Zorin, D.; Bruna, J.; Panozzo, D. Deep Geometric Prior for Surface Reconstruction. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 9 January 2019. [[CrossRef](#)]
12. Maglo, A.; Lavoué, G.; Dupont, F.; Hudelot, C. 3D Mesh Compression: Survey, Comparisons, and Emerging Trends. *ACM Comput. Surv.* **2015**, *47*, 1–41. [[CrossRef](#)]
13. Luebke, D.P. A Developer's Survey of Polygonal Simplification Algorithms. *IEEE Comput. Graph. Appl.* **2001**, *21*, 24–35. [[CrossRef](#)]
14. Bernardini, F.; Mittleman, J.; Rushmeier, H.; Silva, C.; Taubin, G. The Ball-Pivoting Algorithm for Surface Reconstruction. *IEEE Trans. Vis. Comput. Graph.* **1999**, *5*, 349–359. [[CrossRef](#)]
15. Dovrat, O.; Lang, I.; Avidan, S. Learning to Sample. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019. [[CrossRef](#)]
16. Nezhadarya, E.; Taghavi, E.; Razani, R.; Liu, B.; Luo, J. Adaptive Hierarchical down-Sampling for Point Cloud Classification. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 5 August 2020. [[CrossRef](#)]
17. Gong, M.; Zhang, Z.; Zeng, D. A New Simplification Algorithm for Scattered Point Clouds with Feature Preservation. *Symmetry* **2021**, *13*, 399. [[CrossRef](#)]
18. Zhang, K.; Qiao, S.; Wang, X.; Yang, Y.; Zhang, Y. Feature-Preserved Point Cloud Simplification Based on Natural Quadric Shape Models. *Appl. Sci.* **2019**, *9*, 2130. [[CrossRef](#)]
19. Martin, R.R.; Stroud, I.A.; Marshall, A.D. Data reduction for reverse engineering. *Proc. Inf. Geometers Conf.* **1997**, *10*, 85–100.
20. Xu, X.; Li, K.; Ma, Y.; Geng, G.; Wang, J.; Zhou, M.; Cao, X. Feature-Preserving Simplification Framework for 3D Point Cloud. *Sci. Rep.* **2022**, *12*, 9450. [[CrossRef](#)]
21. Leal, E.; Sanchez-Torres, G.; Branch-Bedoya, J.W.; Abad, F.; Leal, N. A Saliency-Based Sparse Representation Method for Point Cloud Simplification. *Sensors* **2021**, *21*, 4279. [[CrossRef](#)]
22. Fan, L.; Atkinson, P.M. An Iterative Coarse-to-Fine Sub-Sampling Method for Density Reduction of Terrain Point Clouds. *Remote Sens.* **2019**, *11*, 947. [[CrossRef](#)]
23. Chen, C.; Yan, C.; Cao, X.; Guo, J.; Dai, H. A Greedy-Based Multiquadric Method for LiDAR-Derived Ground Data Reduction. *ISPRS J. Photogramm. Remote Sens.* **2015**, *102*, 110–121. [[CrossRef](#)]
24. Liu, X.; Zhang, Z. Effects of LiDAR Data Reduction and Breaklines on the Accuracy of Digital Elevation Model. *Surv. Rev.* **2011**, *43*, 614–628. [[CrossRef](#)]
25. Chen, D.; Zhang, L.; Mathiopoulos, P.T.; Huang, X. A Methodology for Automated Segmentation and Reconstruction of Urban 3-D Buildings from ALS Point Clouds. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 4199–4217. [[CrossRef](#)]
26. Honti, R.; Erdélyi, J.; Kopáček, A. Semi-Automated Segmentation of Geometric Shapes from Point Clouds. *Remote Sens.* **2022**, *14*, 4591. [[CrossRef](#)]
27. Albano, R. Investigation on Roof Segmentation for 3D Building Reconstruction from Aerial LIDAR Point Clouds. *Appl. Sci.* **2019**, *9*, 4674. [[CrossRef](#)]
28. Pingel, T.J.; Clarke, K.C.; McBride, W.A. An Improved Simple Morphological Filter for the Terrain Classification of Airborne LIDAR Data. *ISPRS J. Photogramm. Remote Sens.* **2013**, *77*, 21–30. [[CrossRef](#)]
29. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sens.* **2016**, *8*, 501. [[CrossRef](#)]
30. Štroner, M.; Urban, R.; Lidmila, M.; Kolář, V.; Křemen, T. Vegetation Filtering of a Steep Rugged Terrain: The Performance of Standard Algorithms and a Newly Proposed Workflow on an Example of a Railway Ledge. *Remote Sens.* **2021**, *13*, 3050. [[CrossRef](#)]
31. Braun, J.; Braunová, H.; Suk, T.; Michal, O.; Pěťovský, P.; Kurič, I. Structural and Geometrical Vegetation Filtering—Case Study on Mining Area Point Cloud Acquired by UAV Lidar. *Acta Montan. Slovaca* **2022**, *26*, 661–674. [[CrossRef](#)]