3D Cartographic Generalization of LiDAR Point Clouds Based on the Principle of Self-Similarity of a Deterministic Fractal Structure

Generalização Cartográfica 3D de Nuvens de Pontos LiDAR Baseada no Princípio da Autossimilaridade de uma Estrutura Fractal Determinística

Fabiano Peixoto Freiman[1] , Daniel Rodrigues dos Santos[2]
& Allan Rodrigo Nunho dos Reis[3]

[1]Universidade Federal da Bahia, Escola Politécnica, Departamento de Engenharia de Transportes e Geodésia, Salvador, BA, Brasil
[2]Instituto Militar de Engenharia, Secção de Cartografia, Rio de Janeiro, RJ, Brasil
[3]Universidade Federal do Paraná, Departamento de Engenharia Florestal, Curitiba, PR, Brasil
 E-mails: fabiano.freiman@ufba.br; daniel.rodrigues@ime.eb.br; allan.nunho@gmail.com
 Corresponding author: Fabiano Peixoto Freiman; fabiano.freiman@ufba.br

Abstract

The rendering of virtual three-dimensional (3D) structures represented by Point Cloud (PC) allows the representation of internal and/or external environments to buildings. However, the compilation of 3D geometric models is influenced by the intrinsic characteristics of PCs, which can be mitigated by the application of an PC simplification operator. According to the mathematical norms of fractal geometry, it was assumed that a PC is characterized by self-similarity. Two experimental datasets acquired with an SLT in static mode indoors were used. Four tasks were accomplished: sampling and structuring of a PC to solve the problem of random distribution, from an octree structure; estimation of the curvature of the points and the roughness of a neighbourhood for the extraction of edge points by the analysis of self-similarity and application of the Statistical Outliers Remove (SOR) algorithm, for the elimination of outliers points; uniform voxelization, to simplify the intermediate points; application of the Iterative Closest Point (ICP) algorithm to register the sets generated in the same local coordinate system. The use of voxelization was satisfactory, but once the voxel size is manually defined, the PC can be oversimplified and lose essential characteristics. This can be minimized by the primary analysis of the edge points, generating a set that is uniform, less noisy, and self-similar to the original set. To achieve a minimum density of points to model an environment three-dimensionally, one must analyse the geometric self-similarity characteristics of the PC to produce a simplified set self-similar to the original, considering the premises of fractal geometry. It is recommended to create an automatic simplification process to minimize the subjectivity coming from the analyst.

Keywords: Simplification methods; Geometric models; Roughness

Resumo

A renderização de estruturas tridimensionais (3D) virtuais representadas por Nuvem de Pontos (NP) permite a representação de ambientes internos e/ou externos a edificações. Porém, a compilação de modelos geométricos 3D é influenciada pelas características intrínsecas às NPs, o que pode ser mitigado pela aplicação de um operador de simplificação de NPs. Partiu-se do princípio de que uma NP é caracterizada por autossemelhança, conforme as normatizações matemáticas da geometria fractal. Utilizou-se dois conjuntos de dados experimentais adquiridos com um SLT no modo estático em ambientes internos. Foram empregadas quatro tarefas: amostragem e estruturação de uma NP para solucionar o problema da distribuição aleatória, a partir de uma estrutura *octree*; estimativa da curvatura dos pontos e da rugosidade de uma vizinhança para a extração de pontos de borda pela análise da autossimilaridade e aplicação do algoritmo *Statistical Outliers Remove* (SOR), para a eliminação de pontos espúrios; voxelização uniforme, para simplificação dos pontos intermediários; aplicação do algoritmo *Iterative Closest Point* (ICP) para registrar os conjuntos gerados em um mesmo sistema de coordenadas local. O emprego da voxelização foi satisfatório, mas devido à determinação do tamanho do voxel ser definido de forma manual, a NP pode ser simplificada de forma demasiada e perder características essenciais. Isso pode ser minimizado pela análise primária dos pontos de borda, gerando um conjunto menos ruidoso, uniforme e autossimilar ao conjunto original. Para alcançar uma densidade mínima de pontos para modelar tridimensionalmente um ambiente deve-se analisar características de autossimilaridade geométrica da NP para produzir um conjunto simplificado autossemelhante ao original, considerando as premissas da geometria fractal. Se recomenda a confecção de um processo de simplificação automático visando minimizar a subjetividade proveniente do analista.

Palavras-chave: Métodos de simplificação; Modelos geométricos; Rugosidade

1 Introduction

Technological advances and the design of new 3D geospatial data acquisition sensors have made it possible to render virtual three-dimensional (3D) structures (Benita et al. 2020; Nikoohemat et al. 2020). These technologies allow the design of 3D models of cities in the context of Smart Cities (Kumar et al. 2017; Neuville et al. 2018) and Digital Twins (Benita et al. 2020; Döllner 2020; Lehner & Dorffner 2020), to represent internal and/or external environments (Nikoohemat et al. 2020). Among the alternatives for representing content in virtual reality, Point Cloud (PC) gains popularity due to its efficiency in capturing data in loco and encoding it into 3D renderable structures in computing environments.

However, the compilation of 3D geometric models comes up against a recurring problem: the intrinsic characteristics of PCs. When dense, they can portray an environment in three-dimensions. However, they constitute massive, purely geometric data, which presents incompleteness, due to areas of occlusions, and randomness in the distribution of points (Döllner 2020). The central problem consists of the redundancy of a raw PC that, depending on the size of the profiled area and the frequency of profiling of the sensor used, the memory and disk storage requirements can make additional processing impossible without the application of 3D cartographic generalization methods, which are not yet fully available (Sester 2020).

The application of a PC simplification operator represents an alternative to reduce this problem, basically consisting of the process of minimizing the set of points, maintaining the mathematical and statistical characteristics according to the original set, that is, the self-similarity between the original set and its derivatives. Technically, this means that the simplification process must present invariance in its form as the Level of Detail (LoD) changes, keeping its structure identical to the original, falling in the context of a structure fractal (Mandelbroi 1975).

The term fractal is generally applied to diverse constructions, both in the so-called abstract forms and those inherent to nature (Edgar 2008) and are objects of study in the fields of Mathematics and Physics, as laws of formation and scale (Edgar 2008; Mandelbroi 1975). In the context of a PC, the set of points that geometrically defines it must present mathematical self-similarity according to a deterministic fractal structure (Mandelbroi 1975), that is, the mathematical characteristics must be independent of the density of points of a PC. Therefore, the spatial configuration of point distribution must present exact self-similarity, that is, the geometric configurations related to

the original set must be repeated in successive simplified configurations. The great challenge lies in defining a threshold between the loss of geometric characteristics and the minimization of the redundancy of a PC.

Thus, this research started from the principle that a PC, regardless of the profiled environment, is characterized by self-similarity, according to the mathematical norms of fractal geometry. Therefore, from the partitioning of a PC, a simplification operator of the 3D cartographic generalization can be seen as a replica of the whole, in a lower LoD calculated in a recursive procedure, that is, composed of subprocedures that allow the calculation of a fractal dimension that represents the degree of occupation and self-similarity that a structure in space can contain.

1.1 Related Research

Recent advances in 3D data acquisition technologies, such as Terrestrial Light Detection and Ranging (LiDAR) (SLT) sensors, enable the acquisition of PCs with high density per m². However, increasing details usually introduce computational costs in terms of processing and viewing operations. 3D simplification methods allow the minimization of data complexity while maintaining the description of relevant structures. This allows the minimization of problems in the processing of PCs that can become impractical on a large-scale dataset. In particular, there are three ways to simplify PC: A) Mesh-based; B) Based on direct optimization, and; C) Based on sampling; as described in the following subtopics:

– **Mesh-based methods**: It is the preliminary adjustment of a PC to a surface, which undergoes successive simplification processes (Garland & Heckbert 1997). They are iterative methods, present inefficiency in profiling large portions and can be divided into two groups: based on decimation (Asgharian & Ebrahimnezhad 2020) or collapse (Hinderink, Mandad & Campen 2022). The first iteratively removes vertices and then performs a retriangulation process, while the second employs approaches based on edge blending to estimate an optimal position for a vertex (Asgharian & Ebrahimnezhad 2020). These methods assign a weightrelated to the importance of each point in the representation of a surface S based on statistical attributes, such as normal vectors (Hinderink, Mandad & Campen 2022) or the curvature of the point cloud (Asgharian & Ebrahimnezhad 2020). In addition, the determination of an energy function is commonly used to extract edge points (Hinderink, Mandad & Campen 2022).

– **Methods based on direct optimization**: The points of a PC are selected and simplified according to their local mathematical properties (Zou et al. 2020).

The simplification process based on this method assesses the importance of each point in a PC. Normal vectors are often applied and considered the base information to obtain the description of the local geometric characteristics of a neighbourhood of points from iterative strategies. However, although efficient, the algorithms present a high computational cost when applied at larger profiling scales.

– Sampling-based methods: They consist of a point selection scheme whose focus is to maintain the general structure of the object and not its geometric structure, based on the importance of representing the analysed surface. Methods such as the Farthest Point Sampling Histogram (Rusu, Blodow & Beetz 2009), SampleNet (Lang, Manor & Avidan 2020), and S-Net (Dovrat, Lang & Avidan 2019) can be classified as sampling-based methods. The main impasse is related to the non-preserving of the important geometric details of a PC, which can lead to problems in the interpretation and generation of a 3D model.

Regardless of the family of methods, the simplification of PCs inevitably introduces distortions and errors in the rendering of 3D models. The problem lies in the need to develop algorithms that interpret the distribution of the point cloud coherently, imitating a process of human cognition in a computational language. To achieve this objective, this research seeks to select variables and geometric attributes that allow the computational interpretation of the set of points and minimization of their redundancy. The application of mathematical and statistical elements, such as the calculation of normal vectors and the curvature of points belonging to a PC, allows the analysis of the self-similarity of a set P, which is called in this research perceptual metrics.

Motivated by the need to develop 3D generalization operators and by the aforementioned research, it was sought to develop a hybrid process for the simplification of a PC, based on aspects of fractal geometry and perceptual metrics of self-similarity. Therefore, assuming that a PC has mathematical and statistical characteristics (Döllner 2020) and forms a fractal and deterministic structure, to reach a minimum point density condition, a PC composed of a set of points $P = (p_1, p_2, p_{n1-1} \dots p_{n1})$ that undergoes an isotropic simplification process (with the same intensity in all directions), based on the geometric and statistical distribution of a PC is minimized and transformed to the set $P' = (p'_1, p'_2, p'_{n2-1} \dots p'_{n2})$, where $n_2 < n_1$. Consequently, the set P is self-similar to the simplified set P', if P' is invariant after this transformation and is assumed as the main set to model a surface S.

In this context, with the objective of simplifying a PC and minimizing the computational cost the main contributions that this research seeks to provide are: 1) The use of the self-similarity of a fractal structure to specify the sample number of points sufficient to reconstruct a surface S without the need for iterations; 2) Adaptive sampling of edge points based on the roughness calculation of points belonging to a PC in an isotropic process and 3) Preservation of the visual and geometric quality of the dataset in a lower LoD without significant loss of precision.

2 Material and Methods

Two experimental datasets acquired with an SLT in static mode indoors were used, as illustrated in Figure 1.

It was assumed that a PC has mathematical and statistical characteristics (Döllner 2020) and forms a fractal and deterministic structure, to reach a condition of minimum point density. Thus, a PC is composed of a set of points $P = (p_1, p_2, p_{n1-1} \dots p_{n1})$ that undergoes an isotropic simplification process based on the geometric and statistical distribution of a PC is minimized and transformed to the set $P' = (p'_1, p'_2, p'_{n2-1} \dots p'_{n2})$, where $n_2 < n_1$. In this research, geometric characteristics were considered for data minimization and not for topological conditions. Figure 2 shows the methodological flow used to derive P' from a set P.

It is considered that a simplification operator of a PC can be obtained from partitioning the original set into subsets from the establishment of restrictions and analysis of the self-similarity of the neighbourhood of points. The objective is to reduce noise and PC redundancy while preserving edge characteristics and uniformity of the total density of the set. The strategy employed seeks to optimize algorithms in a framework for the development of a point clouds simplification strategy in the context of 3D cartographic generalization, based on statistical self-similarity of the data and calculation of the importance of each point belonging to PC as simplification metrics.

The methodological flow used is subdivided into four tasks (Figure 2). The first consists of the process of sampling and structuring a PC to solve the first problem: random distribution. For this, an octree structure was applied (Rusu & Cousins 2011) that effectively provides a representation and structuring for 3D data. It was chosen to use an octree as the base data structure due to its efficiency in serialization.
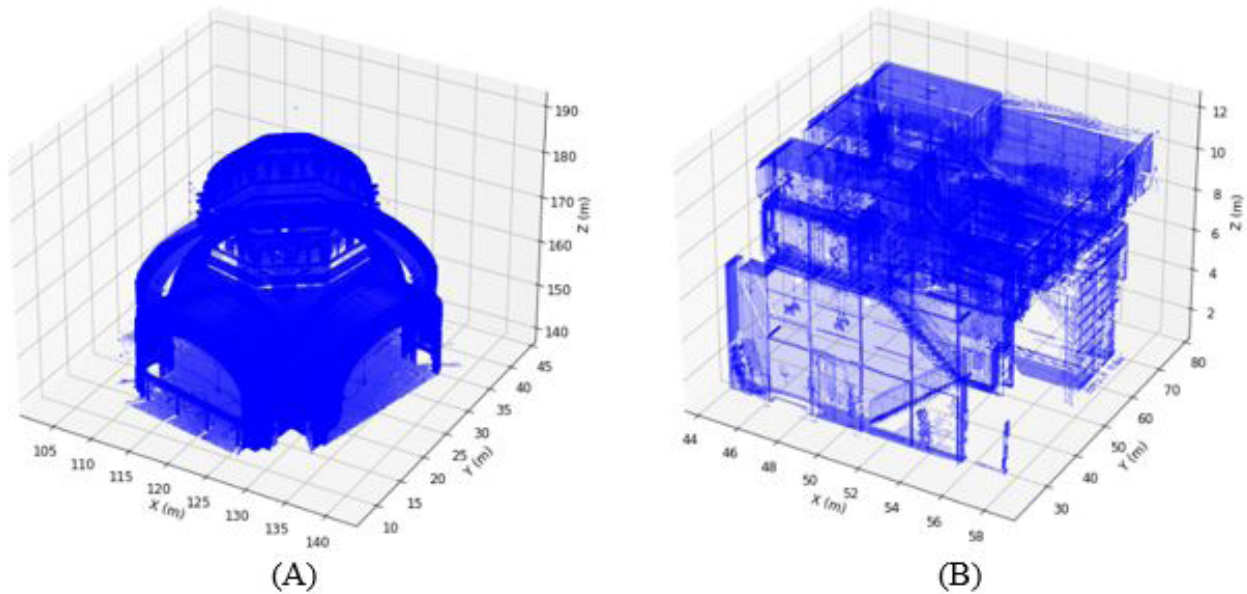
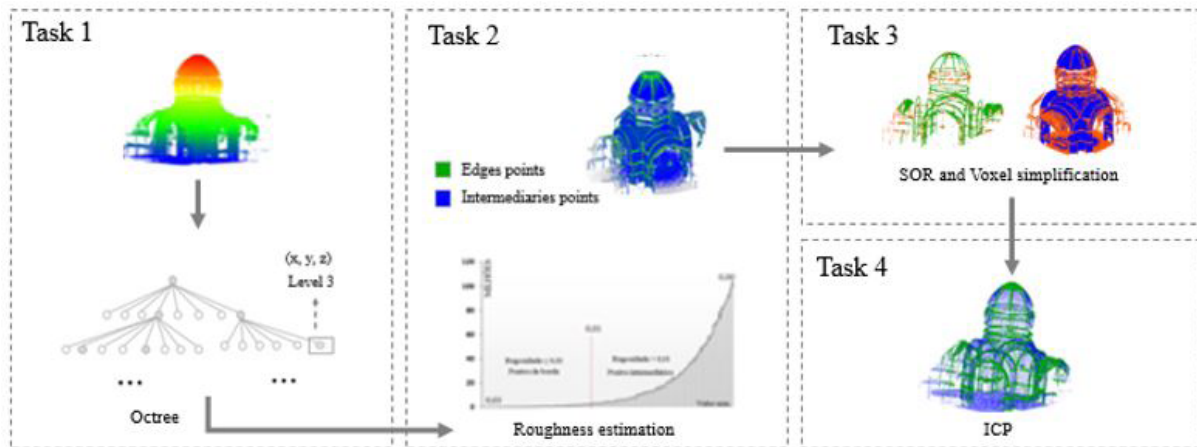**Figure 1** Experimental data: A. Set I; B. Set II.

Figure 2 Application flowchart of the proposed method.

Subsequently, the process of estimating the curvature of points and the roughness of a neighbourhood is applied to extract edge points by the analysis of self-similarity (Pauly, Gross & Kobbelt 2002) and then the algorithm developed by Rusu and Cousins (2011), Statistical Outliers Remove (SOR), is applied to eliminate outliers points in the generated sets. Subsequently, in task 3, the voxelization process based on uniform Lv, Lin and Zhao (2021) is applied to simplify the points classified as intermediate. Finally, the Iterative Closest Point (ICP) algorithm (Besl & McKay 1992) is applied to register the sets generated in the same local coordinate system.

2.1 Extracting edge Points from the Perceptual Roughness Metric

The extraction and simplification of edges of the profiled structures were based on local mathematical metrics of self-similarity, from the analysis of the curvature of the points belonging to the PC, according to Pauly, Gross and Kobbelt (2002). A variance and covariance matrix are an intuitive method for determining the curvature and normal vectors of a PC (Hoppe et al. 1993). Then, considering a neighbourhood N_i around a point $p_i \in R^3$, it is possible to define the covariance matrix presented in Equation 1.

$$C = \begin{bmatrix} p_{i1} - p_1 \\ p_{i2} - p_1 \\ \vdots \\ p_{ik} - p_1 \end{bmatrix}^T \cdot \begin{bmatrix} p_{i1} - p_1 \\ p_{i2} - p_1 \\ \vdots \\ p_{ik} - p_1 \end{bmatrix} \in R^{|N_i| x |N_i|} \qquad \textbf{(1)}$$

From the self-decomposition of C, one can derive the eigenvalues (λ_0, λ_1, λ_2) corresponding by principal components to define an orthogonal reference at the point p_i. According to Hoppe et al. (1993), the eigenvalues with the highest value cover a tangent plane at the point p_i, while those of lesser value are used to approximate the normal surface n_i. Therefore, given the smallest eigenvalue related to p_i of a surface, one can estimate the curvature values (k) of p_i (Pauly, Gross & Kobbelt 2002), according to Equation 2.

$$k(p_i) = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \quad \lambda_0 < \lambda_1 < \lambda_2 \qquad \textbf{(2)}$$

Assuming that $\lambda_0, \lambda_1, \lambda_2$ are the eigenvalues of a local neighbourhood of p_i and v_0, v_1, v_2 correspond to the normal vectors that define a plane $T(x) = (x - p_i)v_0 = 0$, priority is given to the sum of squared distances to neighbours of p_i (Figure 3).

The estimation of the local curvature of p_i is the ideal mathematical variable to analyse the geometric self-similarity and minimization of the redundancy of a set of points. Therefore, using the initial estimate of the curvature of a point p_i, it is possible to estimate the weighted Gaussian mean of the curvatures around a neighbourhood (Equation 3) (An et al. 2021).

$$\hat{k}(p_i) = \frac{\sum_{j \in Ni} k(p_i) \exp\left(-\frac{\|p_j - p_i\|^2}{h} \right)}{\sum_{j \in Ni} \exp\left(-\frac{\|p_j - p_i\|^2}{h} \right)} \qquad \textbf{(3)}$$
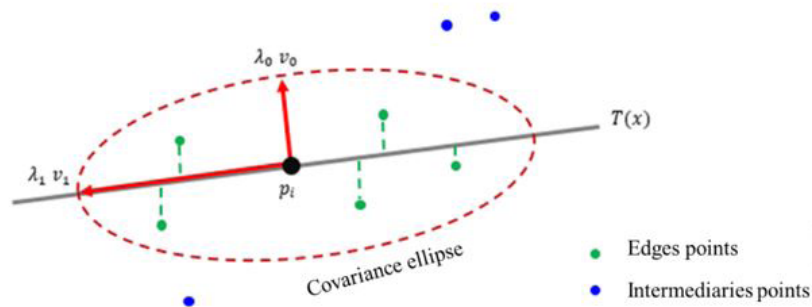
Where h is the search radius of a neighbourhood of p_i. The roughness estimate of a point cloud can be defined as the difference between the curvature (k) and the mean curvature ($\hat{k}$) on point p_i (Rodríguez-Cuenca et al. 2015). The roughness of each point is represented by a simple operator calculated by the variation of the profiled surface. However, for simplification purposes, the output result (segmentation of edges and intermediate points) is still not completely uniform and contains noise that can be minimized. The solution is to resample the data using a simplification and de-noise operator, as presented in the Subsection that follows.

2.2 Voxelization-based Simplification

After extracting the points that allow the identification of edges in the profiled objects, the original PC is subdivided into edge points and intermediate points. It is observed that using only the points belonging to the edges can introduce errors in the processes of defining 3D representations. Therefore, the points classified as intermediate must be associated and used in the previously mentioned set. However, it is necessary to simplify it uniformly over the S surface. For this process, the method developed by Lv, Lin and Zhao (2021) was adapted. The strategy of this step is to subdivide the point cloud into voxel structures (Figure 4).

A voxel is a geometry in a 3D space and corresponds to pixels in a 2D context. From a conceptual point of view, a voxel has a cubic geometry, composed of six faces, eight vertices, and twelve edges. Despite this, the representation is not made according to a polyhedral cube, but from a central point or through the points that represent its vertices (Xu, Tong & Stilla 2021). Therefore, in the context of simplification, a voxel is considered a basic unit that abstracts and structures a space of discrete points, representing a position in a regular cubic grid. In this perspective, the PC classified as an intermediate is subdivided by voxel.
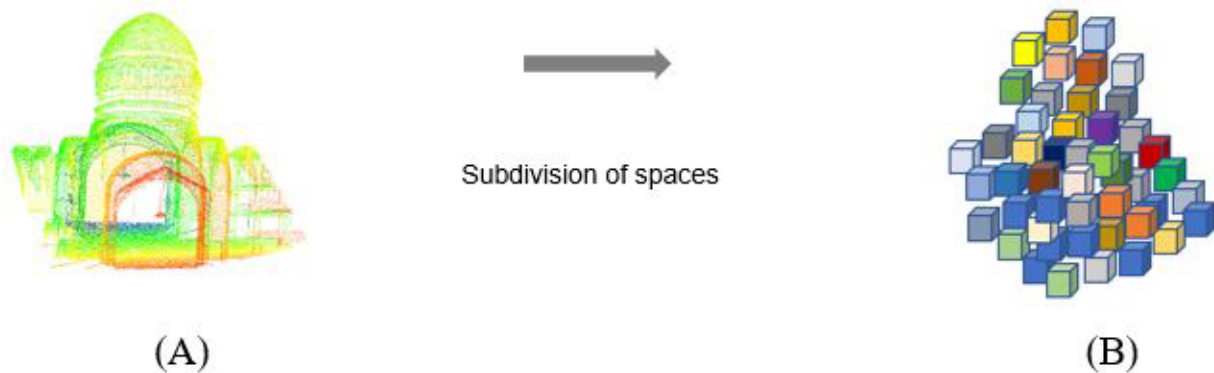

Figure 3 Determination of the curvature of a point.

Figure 4 Point cloud voxelization: A. Original; B. 3D grid (voxel).

The implemented algorithm is performed in three steps: 1) Calculation of an enclosing rectangle for the PC that defines the space to be segmented; 2) The space defined in the first step is subdivided into regularly spaced 3D cuboids of predetermined size (defined by the executor), which become the cells of the voxel structure and; 3) The PC is segmented into small portions from the cuboids, and finally the subdivided set of points is represented by voxels, in which the position and characteristics of the points are calculated from the analysis of the set extracted in each plot. Then, from a set of points $P = (p_1, p_2, p_{n1-1} \dots p_{n1})$ where n is the number of points inside a voxel V_k, and k the voxel index in the voxelized space $V = \{V_1, V_2, \dots V_m\}$, being m the number of all voxels generated.

From this task, the edge points and the intermediate PC are registered using the ICP algorithm to form a generalized cloud ready to serve as a base for the construction of 3D models. To evaluate the effectiveness and the performance of the proposed approach, two validation experiments were used based on data acquired with the SLT from indoor environments. The data processing platform was a Laptop computer with Windows 10, a 1.8 GHz processor and 8 GB of RAM. The Python 3.0 programming language and its libraries, such as the open3D Library (Zhou, Park & Koltun 2018) and Point Cloud Library (PCL) (Rusu & Cousins 2011) were used to estimate the geometric parameters.

3 Results and Discussion

As a first step, it was decided to organize the data sets in the form of an octree, as illustrated in Figure 5. The Octree structure is an efficient way of organizing data, in which a PC is subdivided into nodes (represented by a cube). Each node includes eight secondary nodes (except the leaf, the primary node). The total space is divided into $2^n x 2^n x 2^n$ (Rusu & Cousins 2011), in this case, the cloud is decomposed recursively (one node into eight other subnodes).

This strategy is mainly applied for structuring the PC and minimizing computational costs for future processing steps. With the structured experimental sets, it was possible to determine the roughness of the points that form the PC. From this geometric component, it started the process of analysing the roughness of self-similarity related to the points that compose a sample set.
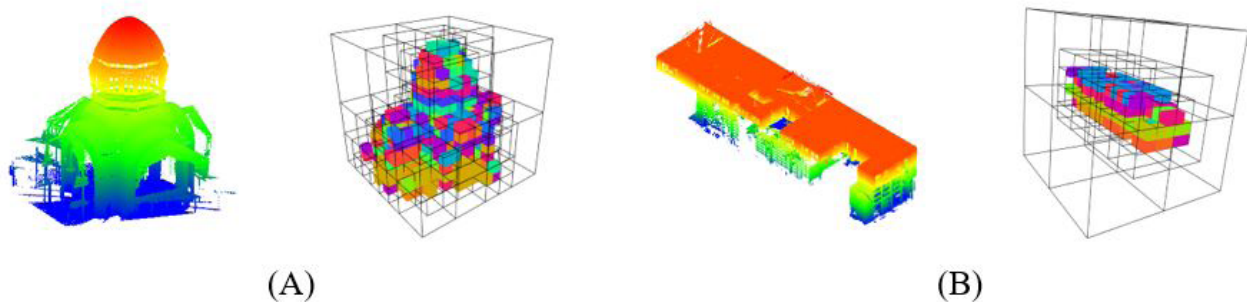

Figure 5 Octree structure: A. Set I; B. Set II.

3.1 Determination of Edge and Intermediate Points

The extraction of points belonging to the edges of the buildings was carried out from the analysis of the roughness of the points. Figure 6 illustrates the number of points in both experimental sets at the corresponding roughness values. In this research, the threshold related to the roughness value was 0.01. This value was defined from plot tests, being the coherent value found for building edges. Points with roughness > 0.01 were classified as edge points (green color – Figure 7). Those with values < 0.01 were categorized as intermediate (blue color – Figure 7).

The analysis of the self-similarity of the geometric roughness parameter proved to be efficient and allowed the extraction of edge points. However, both the edge points and the intermediate sets generated presented outlier points that introduce errors in the geometric modeling process of a surface. To eliminate them and reduce the redundancy of the test PCs, the SRO algorithm was applied. It eliminates noisy data from the sample that does not effectively contribute to the process of building a 3D model.

The profiling of indoor environments was highly complex due to the presence of reflective and translucent surfaces (glass) and occluded areas. Therefore, in the conception of building a 3D model, the step of removing outliers points is relevant to minimize errors in the process of the semantic and geometric interpretation of a PC.

Figure 8 shows the result obtained in the process. The points in red were classified as outliers and eliminated from the sample, while the others were considered as remaining, minimizing the number of points by 10%.

From the segmented PCs and the outliers points eliminated, the simplification process based on voxelization started.

3.2 Voxel-based Simplification

The simplification process based on PC voxelization was applied only to the points classified as intermediate, since they presented the largest number of points that made up the sample and are important in the context of 3D modeling of built environments. One of the main problems in the application of PC simplification processes is defining a specific threshold for reducing the number of points that minimizes the computational cost and maintains minimal geometric properties (explicit representation of the position and topology of a neighborhood of points) stored in abstract 3D structures containing pre-defined positions and attributes so that modeling algorithms can be properly explored (Xu, Tong & Stilla 2021).

Due to this fact, it was decided to perform three specific tests for the sample sets, considering voxel sizes of 0.01, 0.02, and 0.05 m (Figure 9). It is observed that the higher the value of the associated voxel, the lower the number of remaining points. Despite this, too much minimization leads to interpolation errors and estimates of 3D models. Regardless of, it is observed that voxel-based representations are advantageous and efficient in the data compression process.
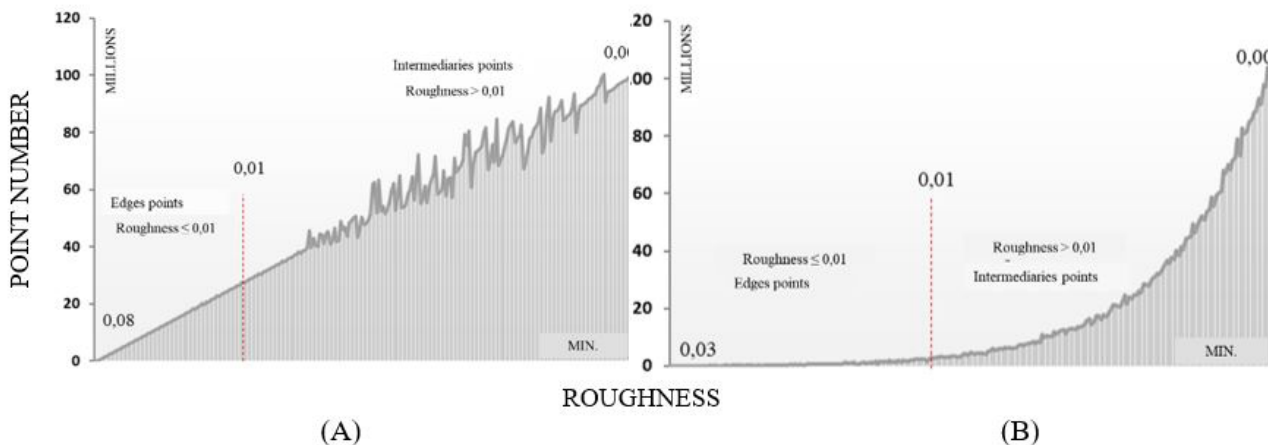

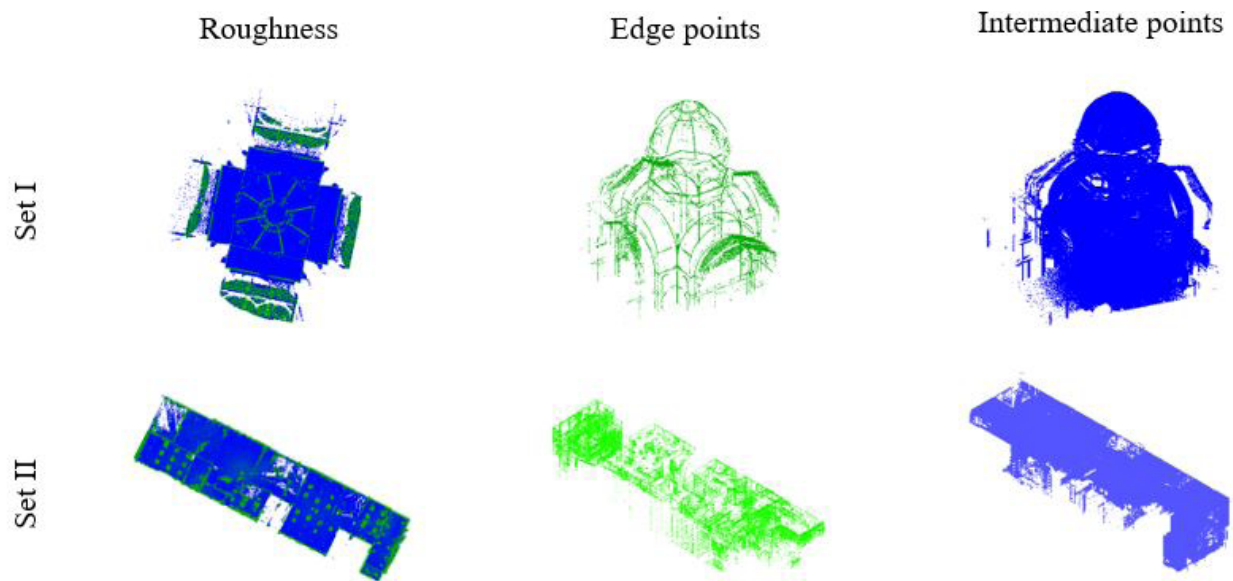**Figure 6** Roughness values: A. Set I; B. Set II.

Figure 7 Extraction of edge and intermediate points: A. Set I; B. Set II.
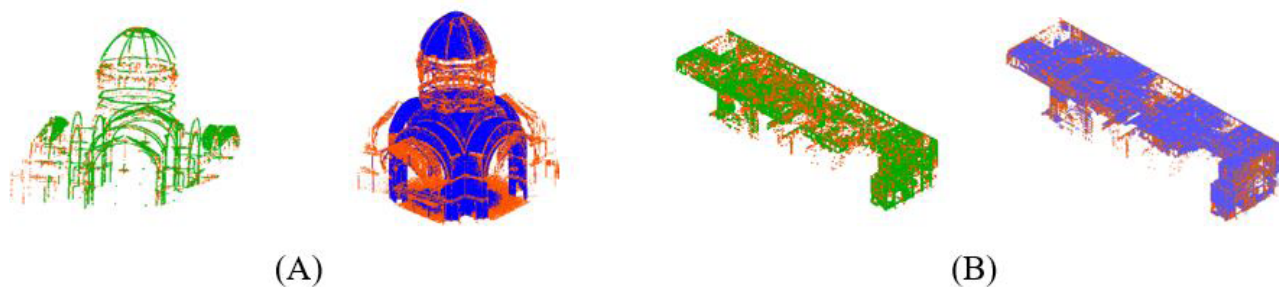

Figure 8 Removal of spirit points: A. Set I; B. Set II.

The objective of the voxel variation is to find an optimal suggestion value for an application in the simplification of PCs, considering the integration of the voxelized set with points categorized as edges in the previous step (Table 1). According to Xu, Tong and Stilla (2021), voxel techniques have a high potential of success to minimize computational cost and, due to their 3D structuring, their application is ideal for monitoring, planning, and navigating tasks in built environments.

For the composition of the final PC, the PC registration process was carried out using a variation of the ICP algorithm between the edge points and the voxelized intermediates, as Figure 10 shows.

The PCs generated were characterized by points concentrated on the edges of the buildings, while on flat (non-rough) surfaces the points were sparse, depending on the size of the voxel used. Table 2 presents the numbers of the remaining points. The original set consists of the point cloud after applying the octree structuring process.

Figure 11 shows the graphic comparison. The addition of edge points in the voxelized data set resulted in an average increase of 23% in the number of points, considering the three tested configurations.

Despite the 23% increase, in both experiments, there was a reduction of 80, 96, and 98%, respectively, for voxels with dimensions of 0.01, 0.03, and 0.05 m. In this case, for the simplification of a PC from profiling of constructions, from the experiments, it was possible to verify that the adoption of the 0.01 m voxel presents adequate simplification rates since a simplification of around 80% is enough to allow processing with lower computational and storage costs, allowing 20% of the original cloud points to remain for future steps.
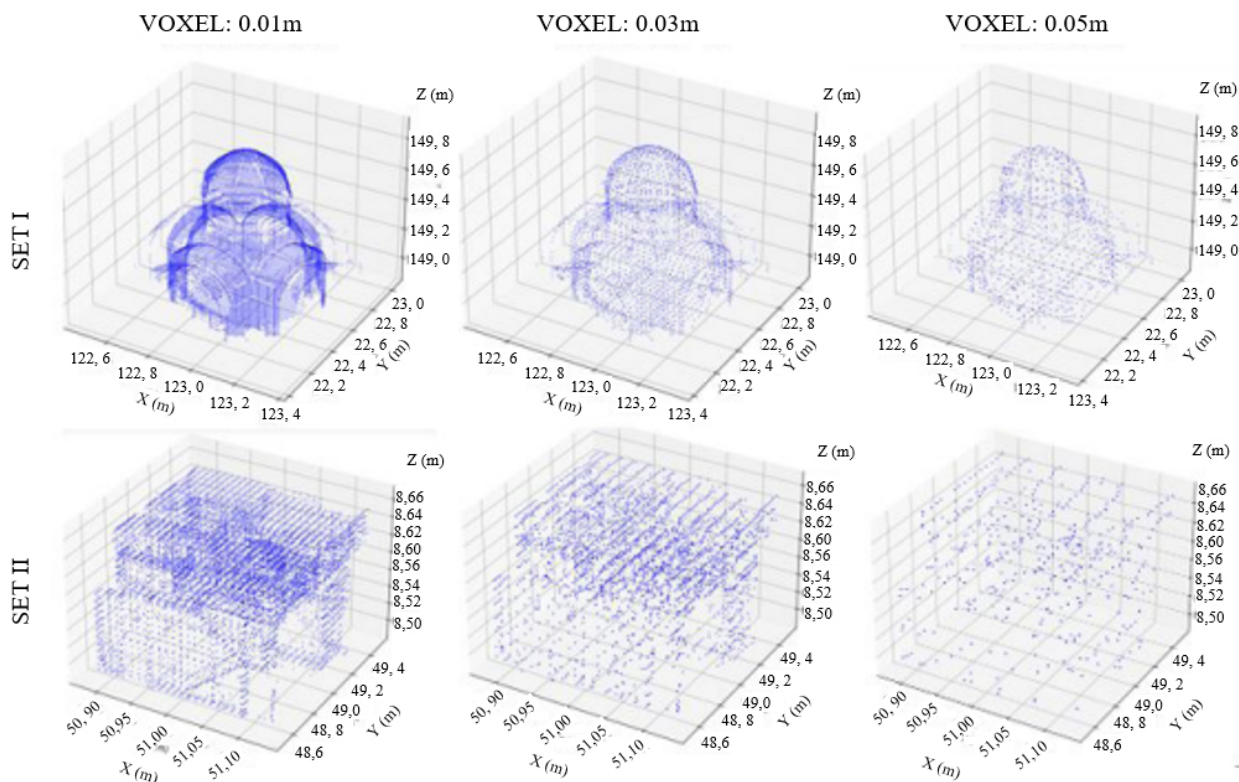
Figure 9 Simplified clouds from voxel: A. Set I; B. Set II.

Table 1 Number of remaining points in the intermediate set.

Experiments	Point numbers			
	Original (Intermediaries)	Voxel 0.01 m	Voxel 0.03 m	Voxel 0.05 m
Set I	4,305,738.00	1,214,465.00	192,063.00	76,504.00
Set II	2,891,867.00	1,033,032.00	135,193.00	53,588.00

The voxels with dimensions of 0.03 and 0.05 m significantly simplified the original sets, remaining less than 5%. This can lead to interpretation errors by computational algorithms in the stages of designing the 3D models. To test the efficiency of the proposal, a comparison was made with two other methods frequently used to simplify PCs.

3.3 Proposed Method Versus Other Simplification Methods

To analyze the experiments, the results obtained with the proposed methodology were compared with two methods used to simplify and minimize the number of points belonging to a cloud. The proposed approach seeks to store a compressed output based on the analysis of roughness values of points belonging to a PC, keeping edge aspects as an essential characteristic of the process. That is, it stores a set of points that coherently represents the changes in perspectives of an environment, minimizing the number of points and consequently the computational cost of processing.

In this case, only the results obtained with the voxel considered ideal (0.01 m dimension) were compared with regularly used algorithms. The Poisson Sample Disk (PSD) and the so-called Random Sample (RS) methods were used, both implemented, respectively, in MeshLab (https://www. meshlab.net/) and CloudCompare (https://www. danielgm. net/cc/). The advantage of the proposed method is the maintenance of edge points automatically, while the other processes minimize the PC without considering that changes in angulations and perspectives can introduce errors in estimates and representations in the making of 3D models.
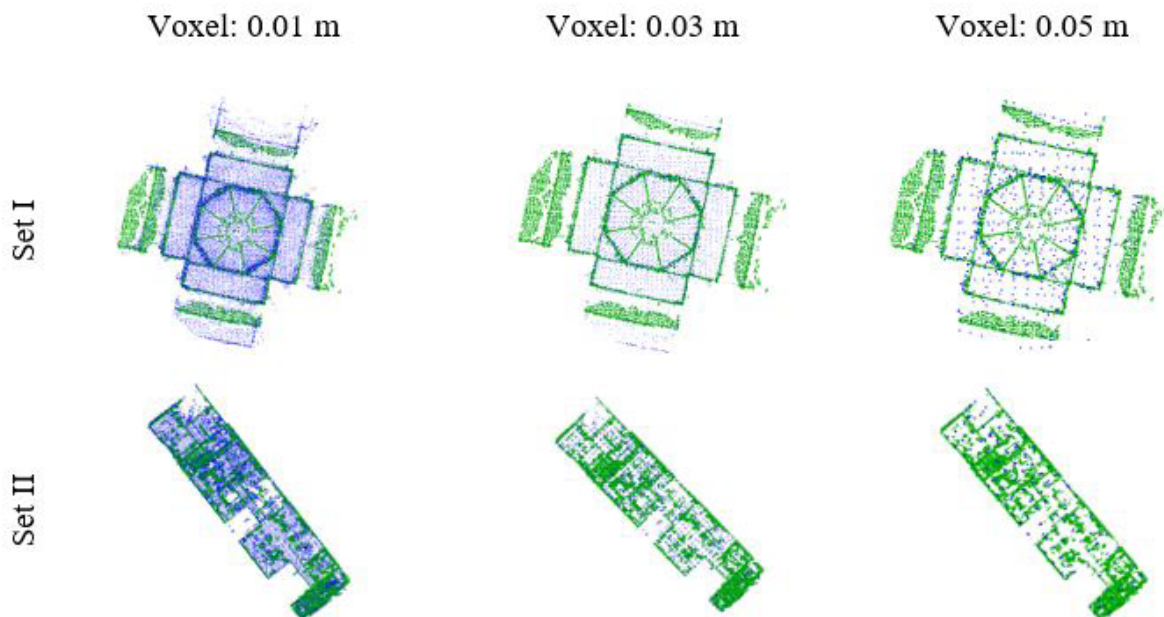
Figure 10 Result of the registration of point clouds: A. Set I; B. Set II.

Table 2 Number of points final clouds.

| Experiments	Point numbers			
	Original	Voxel 0.01 m	Voxel 0.03 m	Voxel 0.05 m
Set I	6,068,202.00	1,224,175.00	221,773.00	106,224.00
Set II	4,982,096.00	1,083,476.00	186,359.00	104,032.00

The comparison test of simplification methods indicated that for both experimental sets, the size of the stored and generated file was larger than the others produced by the application of the tested algorithms. In addition, the geometric quality of the 3D representation of the PC generated with the proposed method was superior. Among the tested algorithms, the RS presented marked limitations, as it is not capable of performing analysis for the definition of the remaining points. The process is based only on the analysis of the distances between the points, leaving only one point as a result of a distance previously established by the analyst.

Mathematically, the results were compared with the definition of the simplification ratio ε. In this case, $\varepsilon = \dfrac{P - P'}{P} * 100$ %, where P is the number of points of the original PC and P' of the simplified sets, as listed in Table 3.

The PSD method allows sampling points based on the principle of proximity by defining the radius of the Poisson disk (user-defined iPCut parameter), which is the half-distance between the two closest samples. In comparison with the RS algorithm, the PSD presents a more uniform distribution over the sampling domain (Hou et al. 2013), however, it does not present the analysis of characteristic points such as the edge ones. The RS strategy allows minimizing the set of points randomly, that is, from the original PC the algorithm minimizes the set to a pre-established number of points. The position of each point does not follow specific rules, but random aspects of distribution. For flat surfaces and with little geometric variability the method can be applied and will be efficient. However, points that are important in the modelling process may be excluded from the sample set.

Despite the simplification ratio being on average 18% lower than the algorithms compared for both experimental sets, there was a simplification rate that allows ensuring that the derived set P' presents adequate geometric self-similarity to the original set P. It is noteworthy that the tests were performed with standard software settings and the parameters can be modified to find a lower simplification rate, but characteristic points, such as edges, will not be preserved.
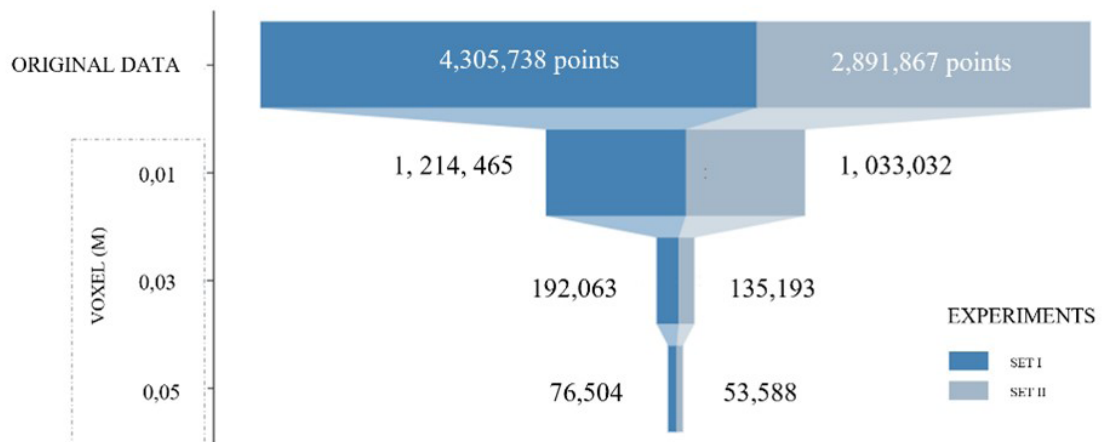
(A)
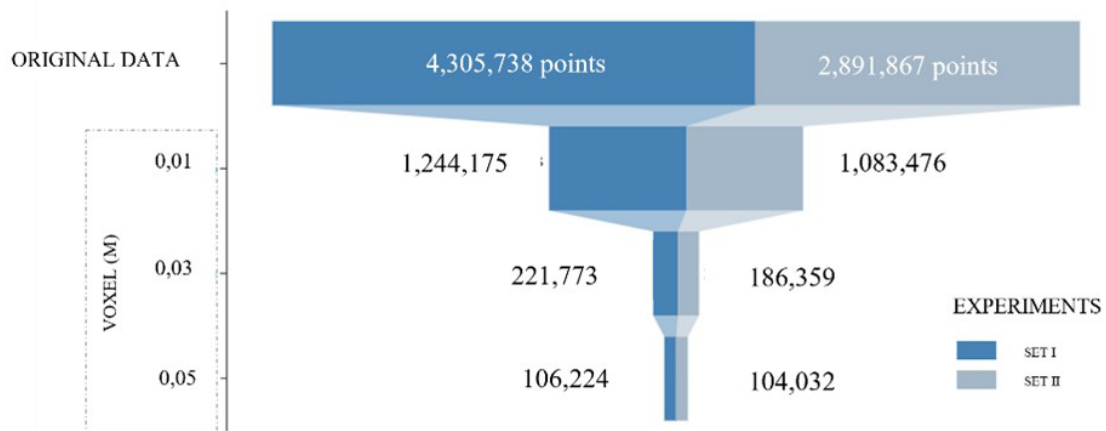
(B)

Figure 11 Simplification process: A. Intermediate (original); B. At the end.

Although the voxelization process requires manual parameterization of the minimum voxel size, integrating edge points into the voxelized PC allows obtaining geometric information from the profiled structures, which brings advantages when compared to the other tested approaches. In principle, the main limitation of the voxel application as a simplification strategy is the manual insertion of its size, which can cause a significant loss of important geometric information for geometric modelling processes of point clouds. The association of the voxelized cloud and edge points allows for minimizing this problem.

The advantage of the procedure is considering the self-similarity of the roughness values of the points belonging to a PC. From this metric perspective, the importance of the points is determined. According to Ji et al. (2019), simplification methods usually sacrifice dataset accuracy to improve processing speed. A proper balance between speed and accuracy must be maintained to minimize errors in future steps. The factors raised by Ji et al. (2019) corroborate the assumptions defended in this research since from the determination of the statistical characteristics of the set of points, there is the formation of a deterministic fractal structure and statistical self-similarity of point distribution, which allows reaching a density condition in an isotropic simplification (voxelization) process to generate a simplified subset that retains the essential geometric features of the original set.

Table 3 Reason for simplification.

Experiments	Reason for simplification		
	Proposal (Voxel 0.01 m)	PSD	RS
Set I	78%	98%	96%
Set II	80%	99%	95%

4 Conclusions

In this work, an approach for the composition of a 3D cartographic generalization operator for simplification of PC using self-similarity based on fractal geometry theories is presented. The approach operates in profiling of outdoor environments to buildings. The method basically explores the estimation of normal vectors and the roughness of point belonging to a PC for inheritance of edge points, which configure as characteristics the self-similarity genes of the set that is associated with a simplified PC. This allows to successfully reduce the number of points os PC and, consequently, to minimize the computational cost of storage and processing.

The use of the voxelization process showed satisfactory results, but due to the manual definition of the voxel size, the PC can be oversimplified and lose essential characteristics. The primary analysis of the edge points allowed to minimize this problem, providing a less noisy, uniform set similar to the original set. Therefore, from the carried-out analyses, it is possible to assume that to achieve a minimum density of points to model an environment three-dimensionally, one must analyse the geometric self-similarity characteristics of a PC to produce a simplified set that is self-similar to the original, considering the premises of fractal geometry. For future researches, it is recommended to create an automatic simplification process to minimize the subjectivity of the analyst.

5 References

An, Y., Wang, L., Ma, R. & Wang, J. 2021, 'Geometric Properties Estimation from Line Point Clouds Using Gaussian-Weighted Discrete Derivatives', *IEEE Transactions on Industrial Electronics*, vol. 68, no. 1, pp. 703–14.

Asgharian, L. & Ebrahimnezhad, H. 2020, 'How many sample points are sufficient for 3D model surface representation and accurate mesh simplification?', *Multimedia Tools and Applications*, vol. 79, no. 39–40, pp. 29595–620, DOI:10.1007/s11042-020-09395-3

Besl, P. & McKay, N. 1992, 'A Method for Registration of 3-D Shapes, Trans', *PAMI*, vol. 14, no. 2, pp. 239-56.

Benita, F., Perhac, J., Tunçer, B., Burkhard, R. & Schubiger, S. 2020, '3D-4D visualization of IoT data from Singapore's National Science Experiment', *Journal of Spatial Science*, vol. 67, no. 1, pp. 157–75, DOI:10.1080/14498596.2020.1726219

Döllner, J. 2020, 'Geospatial Artificial Intelligence: Potentials of Machine Learning for 3D Point Clouds and Geospatial Digital Twins', *Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 88, no. 1, pp. 15–24, DOI:10.1007/s41064-020-00102-3

Dovrat, O., Lang, I. & Avidan, S. 2019, 'Learning to Sample', *Proceedings of the Ieee/CVF Conference On Computer Vision And Pattern Recognition*, pp. 2760–69.

Edgar, G. 2008, *Measure, Topology, and Fractal Geometry*, Springer, New York.

Garland, M. & Heckbert, P.S. 1997, 'Surface simplification using quadric error metrics', *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 1–8.

Hinderink, S., Mandad, M. & Campen, M. 2022, 'Angle-bounded 2D mesh simplification Image 1', *Computer-Aided Geometric Design*, vol. 95, e102085, DOI:10.1016/j.cagd.2022.102085

Hoppe, H., DeRose, T., Duchamp, T., McDonald, J. & Stuetzle, W. 1993, 'Mesh optimization', *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 19–26, DOI:10.1145/166117.166119

Hou, W., Zhang, X., Li, X., Lai, X. & Ding, M. 2013, 'Poisson disk sampling in geodesic metric for DEM simplification', *International Journal of Applied Earth Observation and Geoinformation*, vol. 23, pp. 264–72, DOI:10.1016/j.jag.2012.09.008

Ji, C., Li, Y., Fan, J. & Lan, S. 2019, 'A Novel Simplification Method for 3D Geometric Point Cloud Based on the Importance of Point', *IEEE Access*, vol. 7, pp. 129029-42.

Kumar, K., Ledoux, H., Commandeur, T.J.F. & Stoter, J.E. 2017, 'Modelling urban noise in CITYGML ADE: case of the Netherlands', *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 73–81, DOI:10.5194/isprs-annals-IV-4-W5-73-2017

Lang, I., Manor, A. & Avidan, S. 2020, 'Samplenet: differentiable point cloud sampling', *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7578–88.

Lehner, H. & Dorffner, L. 2020, 'Digital geoTwin Vienna: Towards a Digital Twin City as Geodata Hub', *PFG – Journal of Photogrammetry Remote Sensing and Geoinformation Science*, vol. 88, pp. 63–75, DOI:10.1007/s41064-020-00101-4

Lv, C., Lin, W. & Zhao, B. 2021, 'Approximate Intrinsic Voxel Structure for Point Cloud Simplification', *IEEE Transactions on Image Processing*, vol. 30, pp. 7241–55.

Mandelbroi, B.B. 1975, *The Fractal Geometry of Nature*, WH Freeman and Company, New York.

Neuville, R., Pouliot, J., Poux, F., De Ridder, L. & Billen, R. 2018, 'A Formalized 3D Geovisualization Illustrated to Selectivity Purpose of Virtual 3D City Model', *ISPRS International Journal of Geo-Information*, vol. 7, no. 194, pp. 1–24, DOI:10.3390/ijgi7050194

Nikoohemat, S., Diakité, A.A., Zlatanova, S. & Vosselman, G. 2020, 'Indoor 3D reconstruction from point clouds for optimal routing in complex buildings to support disaster management', *Automation in Construction*, vol. 113, e103109, DOI:10.1016/j.autcon.2020.103109

Pauly, M., Gross, M. & Kobbelt, L.P. 2002, 'Efficient simplification of point-sampled surfaces', *IEEE Visualization*, pp. 1–8.

Rodríguez-Cuenca, B., García-Cortés, S., Ordóñez, C. & Alonso, M.C. 2015, 'A study of the roughness and curvature in 3D point clouds to extract vertical and horizontal surfaces', *IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2015*, Milan, pp. 4602–05, DOI:10.1109/IGARSS.2015.7326853

Rusu, R.B., Blodow, N. & Beetz, M. 2009, 'Fast Point Feature Histograms (FPFH) for 3D registration', *IEEE International Conference on Robotics and Automation 2009*, pp. 3212–17.

Rusu, R. & Cousins, S. 2011, '3D is here: Point Cloud Library (PCL)', *IEEE International Conference on Robotics and Automation 2011 (ICRA 2011)*, Shangai, China, pp. 1–4, DOI:10.1109/ICRA.2011.5980567

Sester, M. 2020, 'Cartographic generalization', *Journal of Spatial Information Science*, vol. 21, pp. 5–11, DOI:10.5311/JOSIS.2020.21.716

Xu, Y., Tong, X. & Stilla, U. 2021, 'Voxel-based representation of 3D point clouds: methods, applications, and its potential use in the construction industry', *Automation in Construction*, vol. 126, e103675, DOI:10.1016/j.autcon.2021.103675

Zhou, Q.-Y., Park, J. & Koltun, V. 2018, 'Open3D: A Modern Library for 3D Data Processing', *ArXiv*, viewed 26 January 2022, <https://arxiv.org/abs/1801.09847>.

Zou, B., Qiu, H. & Lu, Y. 2020, 'Point Cloud Reduction and Denoising Based on Optimized Downsampling and Bilateral Filtering', *IEEE Access*, vol. 8, pp. 136316–26.

Author contributions
Fabiano Peixoto Freiman: conceptualization; formal analysis; methodology; validation; writing-original draft; writing – review and editing; visualization. **Daniel Rodrigues dos Santos:** conceptualization; supervision. **Allan Rodrigo Nunho dos Reis:** writing-original draft; writing – review and editing.

Conflict of interest
The authors declare no potential conflict of interest.

Data availability statement
All data included in this study are publicly available in the literature.

Editor-in-chief
Dr. Claudine Dereczynski

Associate Editor
Dr. Marcus Vinícius Alves de Carvalho