



Article

Uniform and Competency-Based 3D Keypoint Detection for Coarse Registration of Point Clouds with Homogeneous Structure

Fariborz Ghorbani ^{1,2,*} , Hamid Ebadi ¹, Norbert Pfeifer ² and Amin Sedaghat ³¹ Geomatics Engineering Faculty, K.N. Toosi University of Technology, Tehran 19967-15433, Iran² Department of Geodesy and Geoinformation, Technische Universität Wien, 1040 Vienna, Austria³ Department of Geomatics Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz 51666-16471, Iran

* Correspondence: f.ghorbani@mail.kntu.ac.ir

Abstract: Recent advances in 3D laser scanner technology have provided a large amount of accurate geo-information as point clouds. The methods of machine vision and photogrammetry are used in various applications such as medicine, environmental studies, and cultural heritage. Aerial laser scanners (ALS), terrestrial laser scanners (TLS), mobile mapping laser scanners (MLS), and photogrammetric cameras via image matching are the most important tools for producing point clouds. In most applications, the process of point cloud registration is considered to be a fundamental issue. Due to the high volume of initial point cloud data, 3D keypoint detection has been introduced as an important step in the registration of point clouds. In this step, the initial volume of point clouds is converted into a set of candidate points with high information content. Many methods for 3D keypoint detection have been proposed in machine vision, and most of them were based on thresholding the saliency of points, but less attention had been paid to the spatial distribution and number of extracted points. This poses a challenge in the registration process when dealing with point clouds with a homogeneous structure. As keypoints are selected in areas of structural complexity, it leads to an unbalanced distribution of keypoints and a lower registration quality. This research presents an automated approach for 3D keypoint detection to control the quality, spatial distribution, and the number of keypoints. The proposed method generates a quality criterion by combining 3D local shape features, 3D local self-similarity, and the histogram of normal orientation and provides a competency index. In addition, the Octree structure is applied to control the spatial distribution of the detected 3D keypoints. The proposed method was evaluated for the keypoint-based coarse registration of aerial laser scanner and terrestrial laser scanner data, having both cluttered and homogeneous regions. The obtained results demonstrate the proper performance of the proposed method in the registration of these types of data, and in comparison to the standard algorithms, the registration error was diminished by up to 56%.

Keywords: point cloud registration; 3D keypoint detection; uniform; competence; homogeneous structure

Citation: Ghorbani, F.; Ebadi, H.; Pfeifer, N.; Sedaghat, A. Uniform and Competency-Based 3D Keypoint Detection for Coarse Registration of Point Clouds with Homogeneous Structure. *Remote Sens.* **2022**, *14*, 4099. <https://doi.org/10.3390/rs14164099>

Academic Editor: Sander Oude Elberink

Received: 27 June 2022

Accepted: 17 August 2022

Published: 21 August 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/>).

1. Introduction

Recent advances in 3D laser scanner technology have provided a large amount of accurate information as point clouds. These data are used in various fields such as machine vision [1], medicine [2], forestry and urban vegetation [3], geomorphology and surface roughness [4], etc. Instruments such as aerial laser scanners (ALS), terrestrial laser scanners (TLS), mobile mapping laser scanners (MLS), and approaches using image matching are the essential modes for the generation of point clouds [5]. Typically, different scans from various positions are needed to cover all of an object, indoors and outdoors, in all its dimensions. It is necessary to transfer all scans to a common coordinate system in most

applications. In this regard, the process of point cloud registration is considered a critical and fundamental issue.

The point clouds registration process generally consists of two main parts, coarse and fine registration. Each scan is acquired in its local sensor coordinate system with individual origin and orientation of the axes. The largest amount of difference in rotation and non-concurrence between point clouds is reduced in coarse registration. This step is considered as input for fine registration. On the other hand, the error of registration between two point clouds is minimized in the fine registration. The developed algorithms for fine registration (such as ICP [6] and its developments [7]) deliver the final (accurate) results for the registration of point clouds. However, the main challenge for 3D point cloud registration lies in coarse registration [8].

Due to the very high volume of the primary data, 3D keypoint detection is a step that is performed in coarse registration [9]. In the 3D keypoint detection, the initial volume of point clouds is converted into a set of candidate points with high information content. This process will increase the accuracy and reduce the computational cost of the subsequent processing steps. Three-dimensional keypoint detection can be considered to be an operational step in applications such as 3D objects recognition [1,10], SLAM [11], 3D objects retrieval [12], 3D registration [13], and 3D change detection [14].

Various methods have been proposed for 3D keypoint detection. In 2013, Tombari et al. provided a comprehensive assessment of the performance of 3D detectors. They divided 3D detectors into two categories: fixed-scale detectors and adapted-scale detectors. Fixed-scale detectors work on only one scale in the whole process, while adapted-scale detectors consider one scale for each keypoint. The fixed-scale key detector generally consists of two steps. The first step is pruning, and the second is non-maximal suppression (NMS) [15]. In the pruning stage, some criteria remove the least useful information, and only the effective points are transferred to the next stage. In the NMS stage, other evaluations are performed within a radius r to detect the final keypoints. Fixed-scale 3D keypoints detection methods were suggested by [16–20], and adapted-scale detectors were suggested by [21–24]. The 3D keypoints detection algorithms are often designed in the computer vision field, and their primary purpose typically is to detect prominent and distinct keypoints [18]. However, in most applications of photogrammetry and remote sensing, in addition to the distinctness, the spatial distribution and number of 3D keypoints are also of particular importance because it is necessary to create a uniform accuracy of registration in all regions of point clouds.

Limited research has been conducted to develop 3D keypoint detection algorithms in the registration of point clouds on challenging data. Persad and Armenakis [25] presented a method for the initial correspondence of 3D point clouds generated from various platforms. They extracted 2D keypoints from height maps. Their method is designed for man-made scenes and uses images. In 2011, Weinmann et al. [26] proposed an image-based method for quick and automatic TLS data registration. They first extracted the 2D keypoints using the 2D SIFT algorithm and then mapped them into 3D space using interpolated range information. This method needs to extract information from the images, which is a limitation. In 2017, Petricek T et al. [27] presented a method for coarse registration based on an object-based approach for challenging data. The keypoint detection has used a local maximum to determine the salient points. They decomposed the covariance matrix in a neighborhood of each point and evaluated the obtained eigenvalue in the 3D keypoint detection process. Their method only considered the criterion of the quality of points, and no attention has been paid to other criteria. In 2017, Bueno et al. [28] developed the 3DSIFT detection method using geometric features. For this purpose, geometric features are extracted from the decomposition of the covariance matrix in the vicinity, and they displayed point clouds with these features. Instead of applying the 3DSIFT algorithm to the point cloud, they applied the algorithm to the extracted features of the covariance matrix. This method also only considers the quality of the keypoints. In 2019, Zhu et al. [29] combined 3D point cloud data with thermal infrared imaging using a keypoint detection

and matching process. In this research, the keypoints are extracted from the point clouds, and thermal infrared images and the corresponding keypoints are manually determined. Their final output is a 3D thermal point cloud. Recently, the 3D keypoints detection using deep learning methods has also been proposed [13,30–33]. While these methods are emerging, their need for a large amount of training data is considered a limitation. There is little access to this type of training data, especially for real, i.e., not simulated, data.

Although standard 3D keypoint detection algorithms perform adequately in point clouds registration, these algorithms still face challenges, especially in point clouds with homogeneous structures (“Homogeneous structure” means that a large part of the point cloud structure includes limited changes of normal angles in these neighborhoods. In these areas distinct features can hardly be found). These challenges in registration include (i) the quality of the extracted keypoints (limited by defining only one quality criterion), (ii) a lack of proper spatial distribution of keypoints, and (iii) a lack of control over the number of extracted keypoints. These problems prevent achieving a uniform accuracy of registration for the entire point clouds. The main contribution of this paper is to provide a fully automated method for extracting uniform and competency-based keypoints for the coarse registration of point clouds with homogeneous structures. The proposed method effectively provides keypoints to control the quality, spatial distribution, and the number of points necessary to perform the point cloud coarse registration. The benefit of such an endeavor lies firstly in a reliable coarse registration for a wider range of given point clouds and secondly in a more accurate coarse registration, which reduces the number of iterations in fine registration (such as ICP).

The remainder of this paper is organized as follows: in the Section 2, related works are presented, and in the Section 3, the proposed method is explained for 3D keypoints detection. The Section 4 includes the introduction of the data and results, and the Section 5 presents the conclusions and suggestions.

2. Related Works

This section will provide a brief introduction to the standard 3DSIFT and 3DISS keypoint detection algorithms. The following describes the challenges related to these algorithms.

2.1. Keypoint Detectors 3DSIFT and 3DISS

In 3D keypoint detection algorithms, the 3D Intrinsic shape signatures (3DISS) detector [17] and 3D scale-invariant feature transform (3DSIFT) detector [24,34,35] are the most popular algorithms. The SIFT algorithm is a computer machine algorithm to detect and describe local features [36]. This algorithm approximates the Laplacian of Gaussian (LoG) filter using the deference of Guassian (DoG) operator and detects local features with specific dimensions. Inspired by 2DSIFT, 3DSIFT has been developed to detect 3D local features [24]. At first, the scale space is created by the convolution of the Gaussian function with different scale coefficients in the point clouds. The DoG space is then calculated using the difference between two consecutive scales. The keypoints are obtained by calculating local extremes in the DoG space. Finally, the unstable points are refined by the thresholding approach. In this method, only a threshold on quality criterion measures the quality of points and no attention has been paid to the distribution and number of points.

The 3DISS detector was introduced by Zhong et al. in 2009 [17]. In this detector, the quality of points depends on the Eigenvalue Decomposition (EVD) in the covariance matrix obtained from a neighborhood of points (the support region of the points). The covariance matrix for a point is calculated according to the following Equation:

$$\Sigma(p) = \frac{1}{N} \sum_{i=1}^N (q_i - \mu_p)(q_i - \mu_p)^T \quad (1)$$

$$\mu_p = \frac{1}{N} \sum_{i=1}^N q \quad (2)$$

where P is considered a point, N is the number of points that are located in the neighborhood, q_i represents the points in the neighborhood, and μ_p is their average. For the covariance matrix $\Sigma(p)$, the eigenvalues of e_1 , e_2 , and e_3 are sorted in descending order. Points whose ratio of eigenvalues (Equation (3)) is less than a threshold are retained, and other points are removed as unstable points. Among the remaining points, the saliency is determined by the smallest eigenvalue.

$$\begin{aligned} \frac{e_2(p)}{e_1(p)} &< Th_{12} \\ \frac{e_3(p)}{e_2(p)} &< Th_{23} \end{aligned} \quad (3)$$

2.2. Keypoint Descriptor SHOT

In point-based registration approaches, creating key point descriptors is one of the main steps. Among the descriptors presented, the SHOT descriptor [37] is one of the most popular descriptors and has presented appropriate results in comparative research [38]. This descriptor extracts both the geometric and spatial information of points located in the vicinity of keypoints. In the first step, a local reference system (LRF) is generated using eigenvalues obtained from the decomposing of the covariance matrix. Next, along the radial, azimuth, and elevation, the generated LRF is structured into small bins. Histograms are formed from points located in each bin, and the final descriptor is created by connecting all these histograms.

2.3. Limitations of Current Approaches

In this section, the limitations of standard 3D detector methods are discussed. These limitations include (i) the quality of the extracted keypoints (limited by defining only one quality criterion), (ii) a lack of proper spatial distribution of keypoints, and (iii) a lack of control over the number of extracted keypoints. These challenges will be examined in the following.

2.3.1. Measures for Keypoint Quality

The high quality of the points extracted by the 3D detectors will lead to distinct descriptors in the next step. The keypoint quality is limited to one criterion in the standard 3D detectors such as 3DSIFT and 3DISS. It is difficult to achieve appropriate keypoints using only one quality criterion. Quality criteria from different aspects and combining these criteria provides the possibility of achieving more competent keypoints. It is possible to extract various features in the point clouds space. Some of these features have the ability to display point discrimination in a specific neighborhood. Among the most important of these features are the 3D Local Shape Features, 3D Self Similarity, and Histogram of Normal Orientations. In the following, we will introduce these features.

- 3D Local Shape Features

In 2016, Weinmann et al. [39] analyzed 3D environments using feature extraction from point clouds. They selected a suitable neighborhood around the points and extracted geometric features and local shape features using the decomposition of the covariance matrix around each point. In this research, the optimal features are selected using a feature selection approach, and the 3D environment is analyzed using a supervised classification method. They extracted a total of 26 features from the point cloud. Some of the extracted features represent the distinction of points. As the distinction increases, the possibility of producing distinct descriptors in the matching process will rise. The selected distinction features include scattering, omnivariance, anisotropy, and curvature changes. The Equations are:

$$S_\lambda = \frac{e_3}{e_1} \quad (4)$$

$$O_\lambda = \sqrt[3]{\prod_{i=1}^3 e_i} \tag{5}$$

$$A_\lambda = \frac{e_1 - e_3}{e_1} \tag{6}$$

$$C_\lambda = \frac{e_3}{\sum_{i=1}^3 e_i} \tag{7}$$

Parameters S_λ , O_λ , and C_λ , respectively, show the degree of scattering, distribution, and curvature changes in a neighborhood. High values indicate a significant distinction of points. Parameter A_λ represents that a point is located on a continuous surface patch. The lowness of this criterion means the high ability to distinguish points.

- 3D Self Similarity

In 2012, Huang et al. [40] developed the concept of self-similarity on 3D point clouds. Accordingly, self-similarity includes the similarity of normals between a central point and points located in a neighborhood, according to Equation (8). The surface of self-similarity is directly generated by comparing the similarity of center points with neighboring points in a spherical region. According to Figure 1, the more significant the difference between the normal angles is, the lower the self-similarity is. As a result, the criterion considered according to Equation (8) will have a higher value. Conversely, points that have a lower self-similarity will be more salient and distinct. In Equation (8) is the dot product.

$$ss(p, f_{normal}) = \frac{\sum_{i=1}^N \left[\frac{\pi - \cos^{-1}(n(p) \cdot n(q_i))}{\pi} \right]}{N}, \quad ss(p, f_{normal}) \in [0, 1] \tag{8}$$

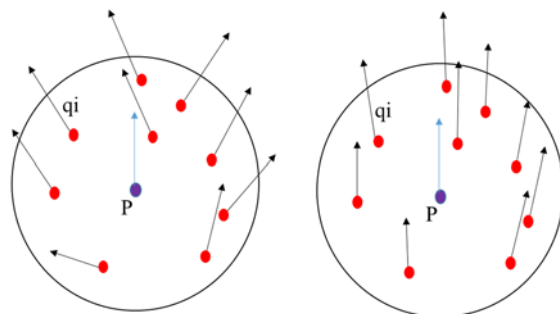


Figure 1. Showing of the normal vector similarity of the center point with other points in a spherical neighborhood.

- Histogram of Normal Orientations

In 2016, Prakhya et al. [41] proposed a keypoint detection method using histograms of normal orientation. They removed the extracted points on flat areas and only selected the located points within the saliency areas. This method does not allow the detection of noises as keypoints. This method considers a spherical neighborhood for each keypoint to produce a histogram of normal orientations. Directional angles are calculated for each neighborhood point according to Equation (9).

$$\theta = \left(\arctan \frac{\|p \times q_i\|}{p \cdot q_i} \right) * \frac{180}{\pi}, \quad \theta \in [0, 180] \tag{9}$$

where $\| \times \|$ represents the cross product and \cdot is the dot product of the keypoint normal vector with the other points in its spherical neighborhood. A direction histogram is formed using the θ values with 18 bins, and the width of each bin is 10 degrees. The histogram is quantified using the normal angles of the neighborhood that fall into each bin. According

to Figure 2, it is easy to see that the histogram values have accumulated in the more homogeneous structure at the first bin, and the other bins will have values close to zero. When the point is located in a cluttered area, most histogram bins will contain values.

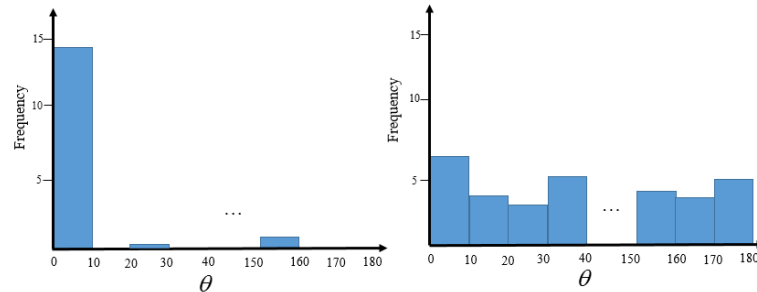


Figure 2. Demonstration of the histogram of normal orientations difference in two homogeneous and cluttered regions.

The kurtosis criterion has been used to consider a quantitative criterion of the separation or peakedness histogram, according to Equation (10).

$$\text{Kurt}(H) = \frac{\sum_{k=1}^N (H_k - \bar{H})^4}{NS_d^4} - 3 \quad (10)$$

$$S_d = \sqrt{\frac{\sum_{k=1}^N (H_k - \bar{H})^2}{N}} \quad (11)$$

In Equation (10), H is the produced histogram, N is the number of bins used (here $N = 18$). H_k represents the value of the k_t bin in the histograms H , \bar{H} is the average of all bins. S_d represents the standard deviation of histogram bins, which is calculated according to Equation (11).

2.3.2. Spatial Distribution and the Number of Keypoints

Another problem with standard 3D detectors includes the lack of control over the spatial distribution of keypoints. In small-scale and standard data, the issue of keypoint spatial distribution generates fewer challenges in the point cloud registration. The role of keypoint distribution is greater in large-scale data. If point cloud data contain an almost homogeneous region with little changes in normal angles and curvature, and other parts contain significant variations of normal angles and curvature, most of the detected points will be concentrated in the cluttered part. This means that registration accuracy is not the same in all regions. The lack of sufficient candidate points in a more homogeneous area increases registration error and, in some cases, causes it to fail. Figure 3 shows the problem of the non-uniform spatial distribution of extracted keypoints by standard detectors.

In this research, the Octree [42] algorithm is utilized to generate an appropriate spatial distribution in keypoint detection. Appropriate spatial distribution refers to the detection of 3D keypoints in the entire point cloud volume. The Octree algorithm has been used in various applications such as 3D segmentation [43], point cloud compression [44], change detection [45], and ICP registration [46]. The Octree is often defined as a tree data structure in which each internal node has exactly eight children, where a 3D space is created by subdivided recursively, dividing it into eight octants [42]. The recursive splitting process is performed only when a cell contains more points than the predefined value. This sequential process continues until it reaches a threshold called depth. The final subdivision leads to eight leaf nodes and stores points in them. Figure 4 shows the division of point cloud space into 3D cells using the Octree method.

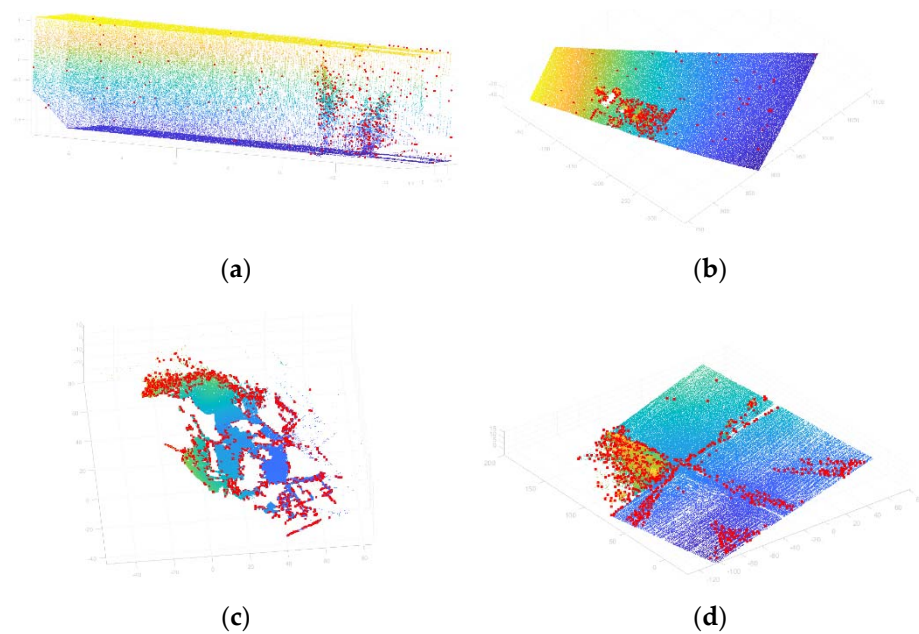


Figure 3. Demonstration of the non-uniform spatial distribution of keypoints ((a–d) are the detected keypoints by 3DSIFT detector).

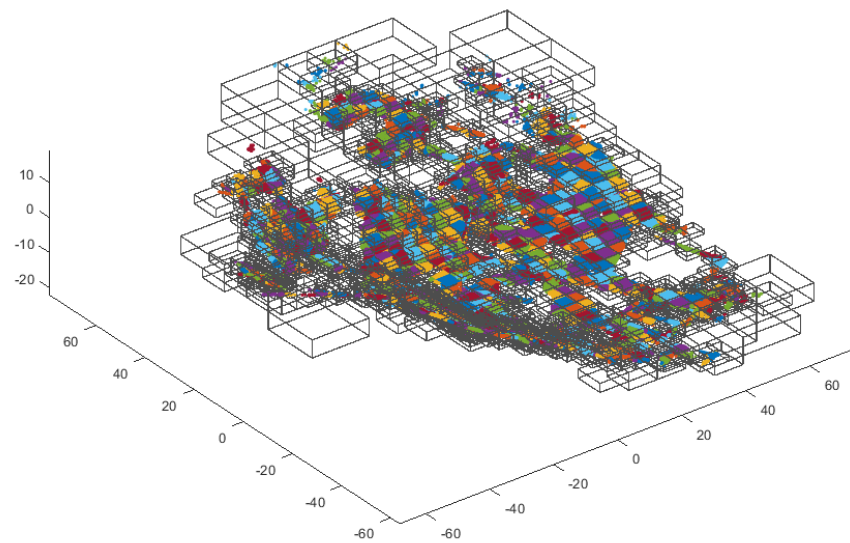


Figure 4. Demonstration of the Octree structure on a point cloud. Data are same with Figure 3d.

Due to the property of the Octree algorithm, the dimensions of each cell in the cluttered areas will be smaller than in homogeneous regions, and there will be more not empty cells there. As a result, the Octree structure supports the extraction of more keypoints in areas with higher discrimination, which are mostly located in cluttered areas. The advantage of having an Octree structure in a cell formation with the same dimensions is that in addition to creating suitable spatial distribution for the extracted keypoints, it provides more chances to more distinctive points in the cluttered area.

The 3DSIFT and 3DISS algorithms detect 3D keypoint only by applying a threshold on the quality criterion. It usually leads to extracting an extra number of keypoints in the regions of the point cloud with cluttered structures. These additional keypoints impose high computational costs in other stages, such as making descriptors, conducting correspondence search, and eliminating wrong correspondences [37]. On the other hand, this number may be much less than the required number in some data with a more homogeneous structure. These problems are mainly due to the nature of the large-scale data, which causes great

sensitivity to the parameters of the detectors, especially the parameter controlling the number of extracted points.

Figure 5 shows the sensitivity of the number of features extracted for some point clouds data. These data include outdoor and indoor scenes, homogeneous areas, and scenes with cluttered structures. This figure presents the number of features extracted by 3DSIFT and 3DISS with different values of their parameters. The T_c parameter for the 3DSIFT detector varies in $[0-0.04]$, and the T_h parameter for the 3DISS detector changes in $[0.5-1]$. This figure shows that there is no optimal value for the thresholding on the parameters of two detectors and displays the high sensitivity of the algorithms to these thresholds. For example, for data A in the 3DSIFT algorithm, keypoints decreases from 24,000 to 2500 when $T_c = 0.001$ changes to $T_c = 0.005$. In the same data, the number of keypoints is increased from 3700 to 12,000 points in the 3DISS algorithm when the change is from $T_h = 0.6$ to $T_h = 0.7$.

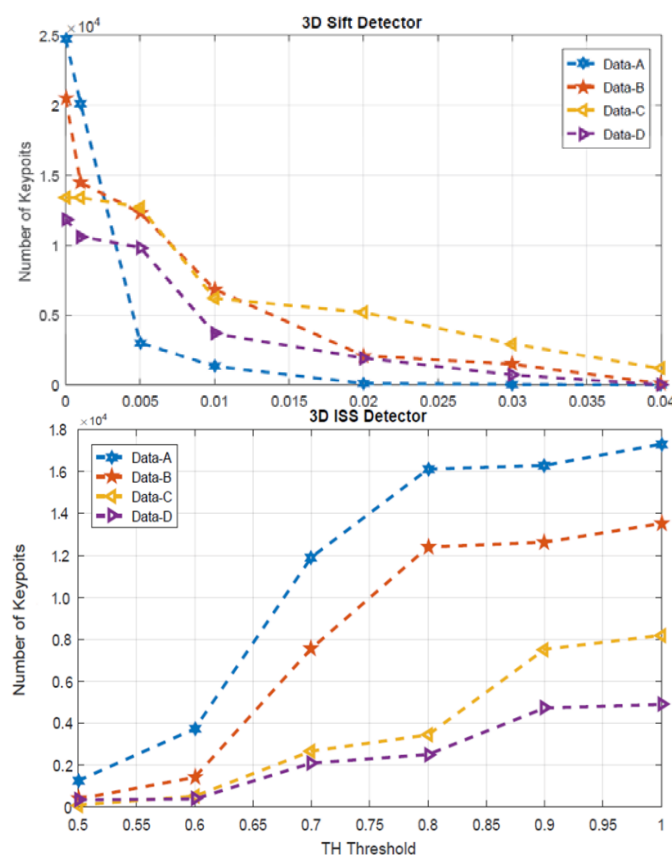


Figure 5. Effects of the T_c and T_h threshold on the number of extracted features in the standard 3DSIFT and 3DISS detectors. Data given in Figure 3.

Research on the selection of high-quality keypoints and appropriate spatial distribution has been conducted on satellite [47,48] and UAV images [49]. Furthermore, it has been performed in 3D space. In 2014, Berretti et al. selected stable key points for person identification using 3D face scans [50]. In 2015 Glira et al. selected optimized points for ALS strip adjustments. Their method is based on the ICP Algorithm and optimized for ALS data [7]. According to the studies, comprehensive research has not been undertaken to detect optimum 3D keypoints to control the quality, spatial distribution, and the number of keypoints in the point cloud registration.

3. Proposed Method

This research presents a uniform and competency-based 3D keypoint detection method for the coarse registration of point clouds with homogeneous structures. Figure 6 shows a

flowchart of the proposed method. At first, the initial keypoints are extracted by 3DSIFT or 3DISS detectors. The keypoint quality is evaluated using different criteria, and the competency criterion is presented to select the keypoints with the highest competence. The Octree structure is then used to create a uniform spatial distribution in the point cloud. In each cell, the keypoints with the highest quality will be extracted. It will be proportional to the number of points assigned to each cell. The keypoints are described using the signature of histograms of orientations (SHOT) descriptor to perform the matching process. Finally, the required evaluation is accomplished after the matching and registration process.

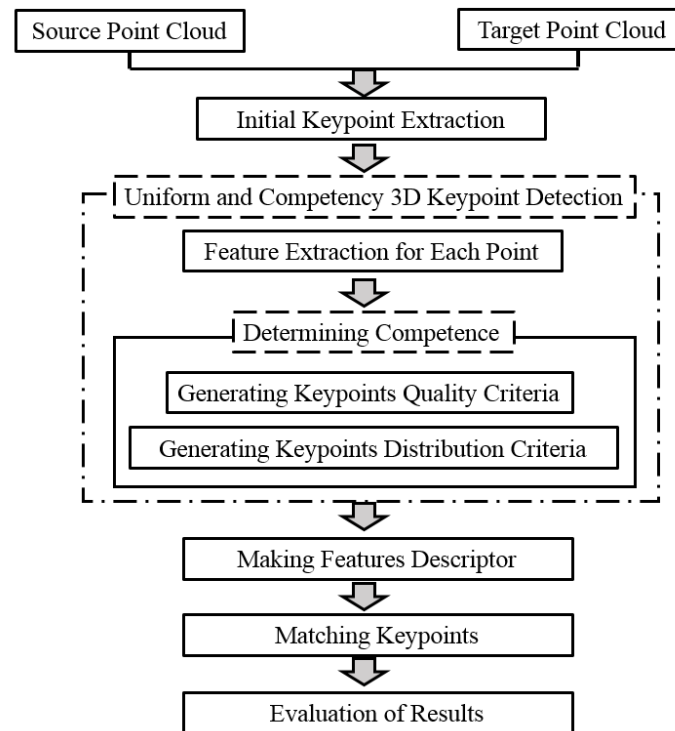


Figure 6. Diagram of Proposed approach.

Following [47], the proposed method has been developed for extracting a sufficient number of 3D keypoints with high quality and in a suitable spatial distribution in the point cloud space. The steps of the algorithm are as follows:

1. Extraction of initial keypoints by detector algorithms (3DSIFT or 3DISS). At this stage, by selecting small thresholds, a relatively large number of points are extracted.
2. Estimate competence for each initial keypoint by:
 - (1) The 3D shape features (Scattering, omnivariance, anisotropy, change of curvature), the 3D self-similarity feature, and the Histogram of Normal Orientations feature are extracted for each initial keypoint. These properties are considered as a vector with n components as follows:

$$X^j = \{x_i^j\}; \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m \quad (12)$$

where i is related to the initial point number and j is for the criteria number (m is equal to 6 here).

- (2) The ranking vector of the initial points is determined for each criterion. The highest value will be the first rank, and the lowest value will be the last rank for all criteria except anisotropy (A_λ). The criterion anisotropy is the opposite,

and the lowest value will have the highest rank. The ranking vector of each criterion is displayed as follows:

$$R^j = \{r_i^j\}; \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m \quad (13)$$

where r_i^j is related to the i_{th} keypoint in the j_{th} criterion vector and indicates the rank of this keypoint against others in that criterion.

- (3) The competence of initial keypoints is calculated using a combination of all criteria as follows:

$$C_i = \sum_{j=1}^m w^j (n - r_i^j) \quad (14)$$

where C_i corresponds to the competence of the j_{th} keypoints, r_i^j indicates the rank of the i_{th} keypoints in the j_{th} criterion, n is equal to the number of keypoints, m is the number of criteria. w^j is the weight parameter for the j_{th} criterion, which is used to control the effect of different criteria on the keypoints quality. The sum of the weights of the various criteria must be equal to one. The higher the C -competence criterion is, the better the complication is, and it has more probabilities of success in the matching process.

3. The control of the keypoint spatial distribution is conducted by cell formation in point clouds based on the Octree structure. The details of this process are as follows:
- (1) The point cloud space is cellularized using the Octree structure. Using the depth parameter, it is determined to what extent the cell formation will continue.
 - (2) The total number of required keypoints (N) is determined for detection. The N parameter controls the number of required keypoints for extraction.
 - (3) The number of extractable keypoints in each cell is calculated according to the value of competence and the number of initial points located in each cell. In fact, in this step, it is determined how many of the total numbers of required keypoints (parameter N) are allocated to each cell. This process is determined as follows:

$$N_{ck} = N \left[\frac{w_n \cdot n_k}{\sum n_k} + \frac{w_c \cdot \bar{C}_k}{\sum \bar{C}_k} \right] \quad (15)$$

where n_k is the number of initial keypoints located in a specific cell and \bar{C}_k represents the average competency of the keypoints in that cell. w_n and w_c are weights related to n_k and \bar{C}_k . In general, the higher the number of initial keypoints in a cell and the higher the average competency of the keypoints, the more points are extracted from that cell.

3.1. Matching and Registration

After producing uniform and competent keypoints by using the proposed method, the next step is applying a local feature descriptor for the 3D keypoints and performing the registration process. Here, the SHOT descriptor is used for describing the 3D keypoints. The registration process is performed using a RANSAC-based approach, following [51]. In this method, according to Figure 7, two-point clouds are considered as the source (P_s) and target (P_t). The closest keypoint in the target point cloud is determined for each of the described keypoints in the source point cloud by calculating the Euclidean distance among their descriptors. Then, these points are considered initial correspondence points if their Euclidean distance is less than a threshold. In order to remove the wrong correspondences, the relationship between two point clouds is determined using a 3D transformation such as 3D similarity. In this method, The M-estimator sample consensus (MSAC) model [52], an extension of RANSAC, is used to estimate 3D similarity transformation parameters. The source point cloud is transferred to the target point cloud using the estimated parameters.

Then the distance between the corresponding points is used as the error of each corresponding pair. The root means square error (RMSE), as a criterium to measure the error of registration, is compared with a threshold limit. If it is greater than this value, the point with the highest error is removed from the registration process. This process continues until reaching an RMSE with the desired value.

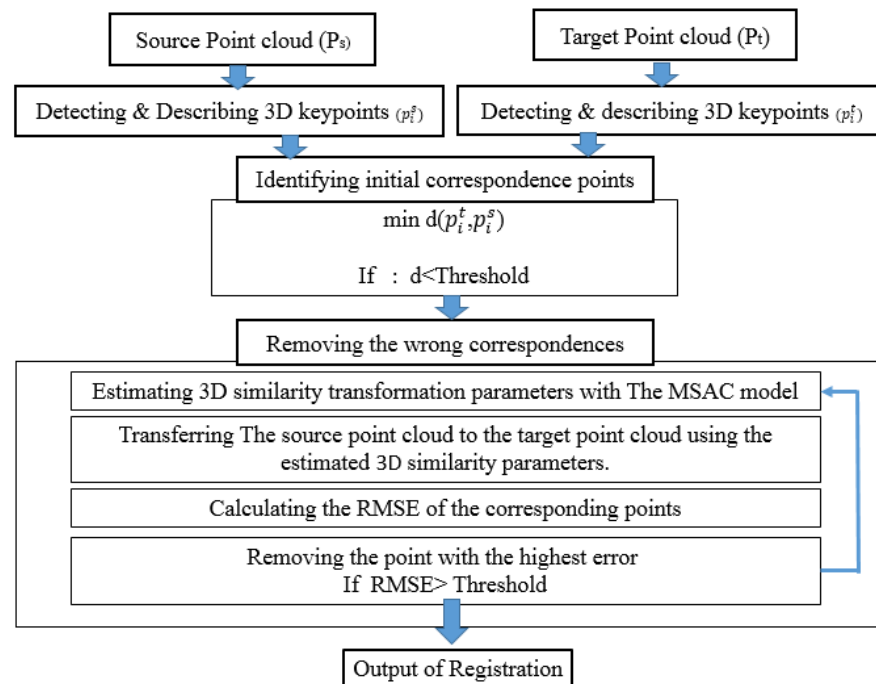


Figure 7. A diagram of the steps of matching and registration.

3.2. Evaluation Criteria

Various criteria have been used to evaluate the proposed method. These criteria include repeatability, recall, precision, spatial distribution, and positional accuracy. Repeatability is a percentage of detected keypoints in two point clouds that correspond to each other after the transfer to the same coordinate system. The recall and precision criteria were calculated using the following Equation:

$$\text{Recall} = \frac{CM}{TM} \quad (16)$$

$$\text{Precision} = \frac{CM}{CM+FM}$$

In Equation (16), CM (Correct Match) and FM (False Match) are the number of correct and false correspondences obtained by the proposed method in the matching results, respectively, and TM (Total Match) is all the correct matches available in the matching process. The geometrical relationship between each pair of point clouds must be known to determine these criteria. In data B , the secondary data were simulated to create registration conditions, so the geometric relationship between them is obvious, and the exact position of the corresponding points can be easily calculated. The 3D similarity transformation was calculated manually to establish a geometric relationship between two point clouds in data A and data C . For this purpose, reflective targets were installed in the locations. An expert operator has manually identified these points in each point cloud. These corresponding points were used to calculate the 3D similarity model. Furthermore, the ground truth transformations are available for the data D , and the user can use them to transfer point clouds coordinate system. This paper has used a spatial threshold equal to $2mr$ (mr is the mean of resolution in point clouds) to distinguish the corresponding points from the non-corresponding points.

The local distribution density and the global coverage index have been used to evaluate the spatial distribution quality of points. These criteria are calculated based on the Voronoi diagram as follows:

$$\delta_{dense} = \frac{\sum_{i=1}^n |V_i - \bar{V}|}{\sum_{i=1}^n V_i}; \quad \alpha_{cov} = \frac{\sum_{i=1}^n V_i}{V_{Total}} \quad (17)$$

where δ_{dens} is the local distribution density and α_{cov} is the global coverage, V_i is the volume of the i_{th} Voronoi cell, \bar{V} is the mean volume of all the Voronoi cells, n is the total number of Voronoi cells and V_{Total} is the total volume of common parts in two-point cloud data. The smaller δ_{dens} and the larger the value of α_{cov} indicates the more appropriate spatial distribution of the corresponding points.

Two criteria have been used to evaluate the positional accuracy of registration. The first criterion is proposed to measure rotational and translation errors. Generally, they are used to validate the point cloud registration [27,53]. Suppose the source point cloud (P_s) is transferred to the target point cloud (P_t) by the transformation function ($T_{s,t}$). The remaining value of transformation ($\Delta T_{s,t}$) is defined as follows:

$$\Delta T_{s,t} = T_{s,t} \left(T_{s,t}^G \right)^{-1} = \begin{bmatrix} \Delta R_{s,t} & \Delta t_{s,t} \\ & 1 \end{bmatrix} \quad (18)$$

where $T_{s,t}$ is the estimated transformation function from P_s to P_t and $T_{s,t}^G$ is the ground truth transformation function. Then the rotation error ($e_{s,t}^r$) and the translation error ($e_{s,t}^t$) from P_s to P_t are calculated based on the rotational components ($\Delta R_{s,t}$) and the translation components ($\Delta T_{s,t}$) as Equation (19). In this Equation, $tr(\Delta R_{s,t})$ represents the trace ($\Delta R_{s,t}$).

$$\begin{cases} e_{s,t}^r = \arccos\left(\frac{tr(\Delta R_{s,t})-1}{2}\right) \\ e_{s,t}^t = \|\Delta t_{s,t}\| \end{cases} \quad (19)$$

The second criterion is The RMSD (The root-mean-square distance) criterion, and it is a combination of two rotational and translation errors. In this criterion, the root-mean-square distance is calculated between the registered point cloud by the algorithm and the point cloud in its true position. RMSD is utilized for assessing positional accuracy. If we consider the source point cloud after the transfer as P and the same point cloud in its true position as G , the RMSD value can be calculated from the following Equation:

$$RMSD = \frac{\sum_{i=0}^n \|p_i - g_i\|_2}{n} \quad (20)$$

where $\|\cdot\|$ is the Euclidean distance, and n is the total number of points. It should be noted that the point clouds are on the same scale.

4. Experiments

4.1. Data

The method proposed in this study was tested and analyzed across four different point clouds. The most important attribute of these data is the existence of large regions of point cloud with an almost homogeneous structure and a cluttered part simultaneously. In the following, we will review these data. In addition, they are shown in Figure 8.

- Data A: These data were obtained by terrestrial laser scanner in an indoor environment at two different stations with 100% coverage. They were taken from a corridor of a university building (TU Wien). In Figure 8a, we can see some objects concentrated in a small area of the corridor, and there are also some installed signs along the corridor.
- Data B: This point cloud was obtained using aerial laser scanners at Ranzbach (a forested area in lower Austria, west of Vienna). It was provided by the company Riegl

and acquired in April 2021. The used scanner in this data is VQ-1560 II-S. The scan area consists of a flat area in the middle and a forest of different tree types around it. The selected area included a small section of dense trees, and other areas had an almost homogeneous surface. The secondary data were simulated to create the co-registration conditions. The simulated data were generated by applying shifts, rotations, and density changes in the point cloud to examine the co-registration process.

- Data C: These data were taken by Terrestrial laser scanners in an outdoor environment at two different stations with 100% coverage. The geographical location of the data is in the flat, rural area of Vienna (22nd district). These data can be divided into three parts. The first part consists of dense trees. The second part consists of flat land, and the third part includes agricultural land with little topographic changes.
- Data D: These data are a subset of the ETH PRS TLS benchmark as the Courtyard dataset. Point clouds in this benchmark were taken by a Z + F Imager 5006i and Faro Focus 3D and provided in an outdoor environment for the TLS point cloud registration. There are no vertical objects in these data. It was generated to create DTMs.

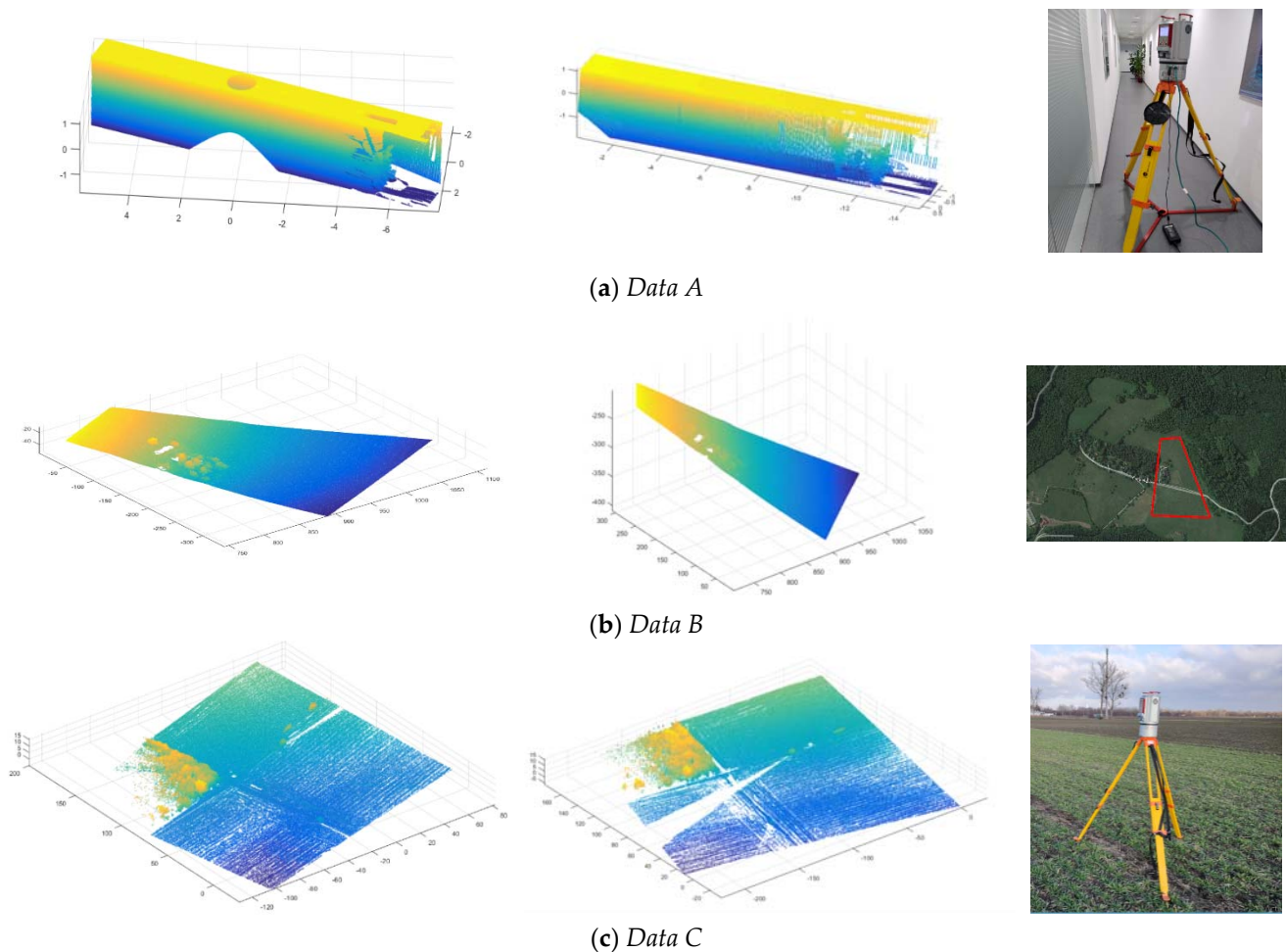
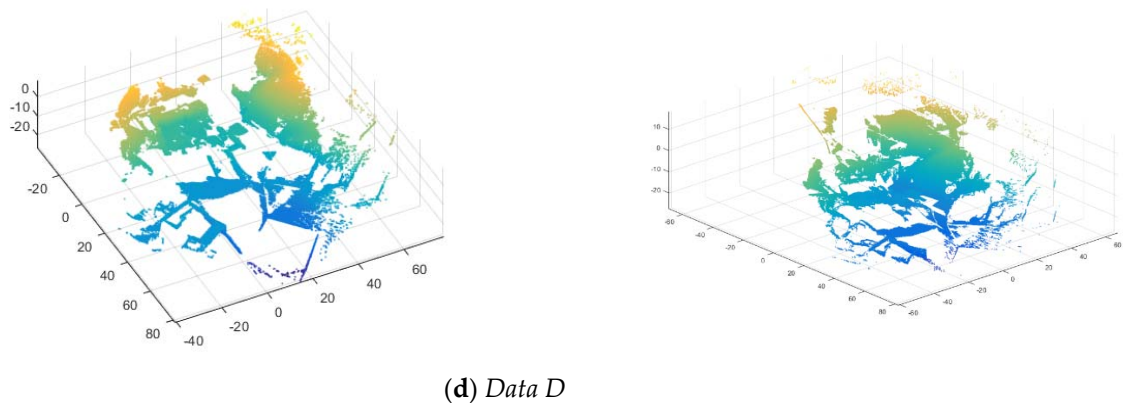


Figure 8. Cont.



(d) Data D

Figure 8. Demonstration of a scan of the data used.

4.2. Results

This section presents the performance of the proposed approach in the point cloud registration. All of the steps of the proposed algorithm were implemented on MATLAB (R2021a). In the evaluation process, two 3DSIFT and 3DISS detectors algorithms and a SHOT descriptor algorithm are used. These algorithms were implemented in PCL [24]. The parameters of the proposed method include the number of points to be extracted (N), the maximum depth in the Octree structure ($O_{C_{Depth}}$), which represents the maximum number of times a bin can be divided and the weights associated with each of the quality criteria (w_1, \dots, w_6), and the weight related to determining the number of points in each bin (w_n and w_c). To determine the optimal weights, different values between $[0, 1]$ were considered for each of the parameters. Then, the optimal value for each of the parameters was determined by implementing the proposed method on all of the data and considering the RMSD criteria. Table 1 shows the considered values for the parameters of the proposed algorithm. All of these optimal values were determined experimentally by applying different amounts of the parameters. Furthermore, the results of uniform and competency 3D keypoints detection by the proposed method (3DUCSIFT and 3DUCISS) are presented in Figure 9.

Table 1. The parameters of the 3D Keypoints detection method.

Parameter Name	Denotation	Selected Value	
Number of required Keypoints	N	1% total of points	
Maximum depth in Octree structure	$O_{C_{Depth}}$	4	
Weights related to competency measure computation	Scattering	W_1	0.1
	Omnivariance	W_2	0.1
	Anisotropy	W_3	0.1
	Change of curvature	W_4	0.1
	Self-Similarity	W_5	0.3
	Histogram of Normal Orientations	W_6	0.3
The weight is related to determining the number of points per cell	Average competency	W_c	0.5
	Number of keypoints	W_n	0.5

The matching results for the proposed method (3DUCSIFT and 3DUCISS) as well as the standard 3DSIFT and 3DISS methods on the considered data are shown in Figure 10. In Table 2, the number of detected keypoints is presented. As mentioned earlier, the proposed method can control the number of detected keypoints. As a result, to perform a proper comparative process, the number of detected keypoints by the proposed algorithm is

considered close to standard algorithms. On the other hand, the mean resolutions (mr) of the used data are 5, 60, 45, and 15 cm, respectively. These different resolutions indicate a variety of densities in the evaluation of the methods.

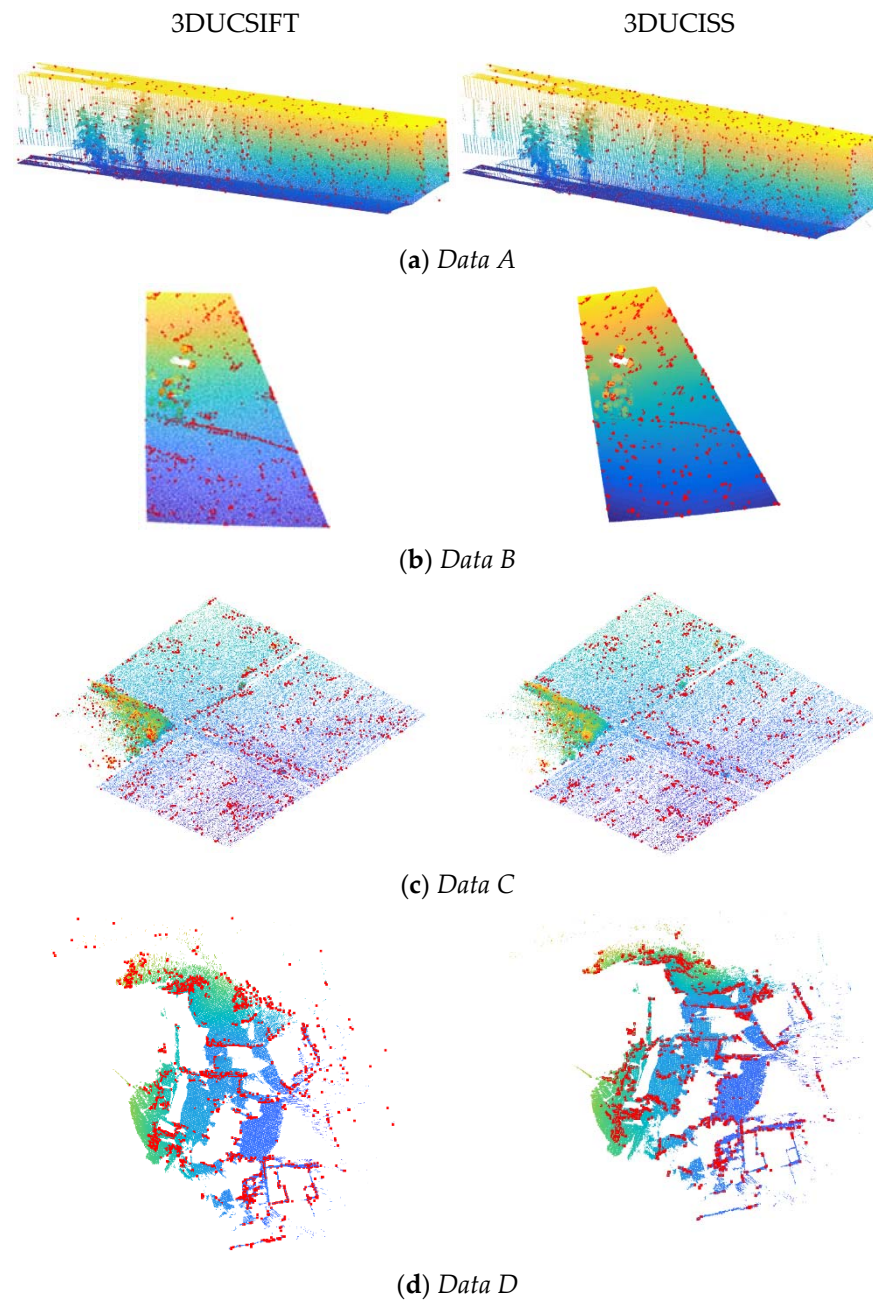


Figure 9. The result of 3D keypoints detection by the proposed method. The left column is the 3D keypoints detection result of the proposed method using 3DSIFT(3DUCSIFT) and the right column is the 3D keypoints detection results of the proposed method using 3DISS (3DUCISS).

Figure 10 shows that in these type of data, the spatial distribution of corresponding points is limited for standard algorithms, and the corresponding points are not detected in areas with homogeneous structures. However, in the proposed method, the spatial distribution of the corresponding points has improved significantly.

The results of different evaluation criteria to compare the performance of the proposed method are presented in Figures 11–13. Figure 11 shows the results of the Recall, Precision,

and Repeatability criteria. Figure 12 presents the spatial distribution results, and Figure 13 shows the positional accuracy results of the methods tested.

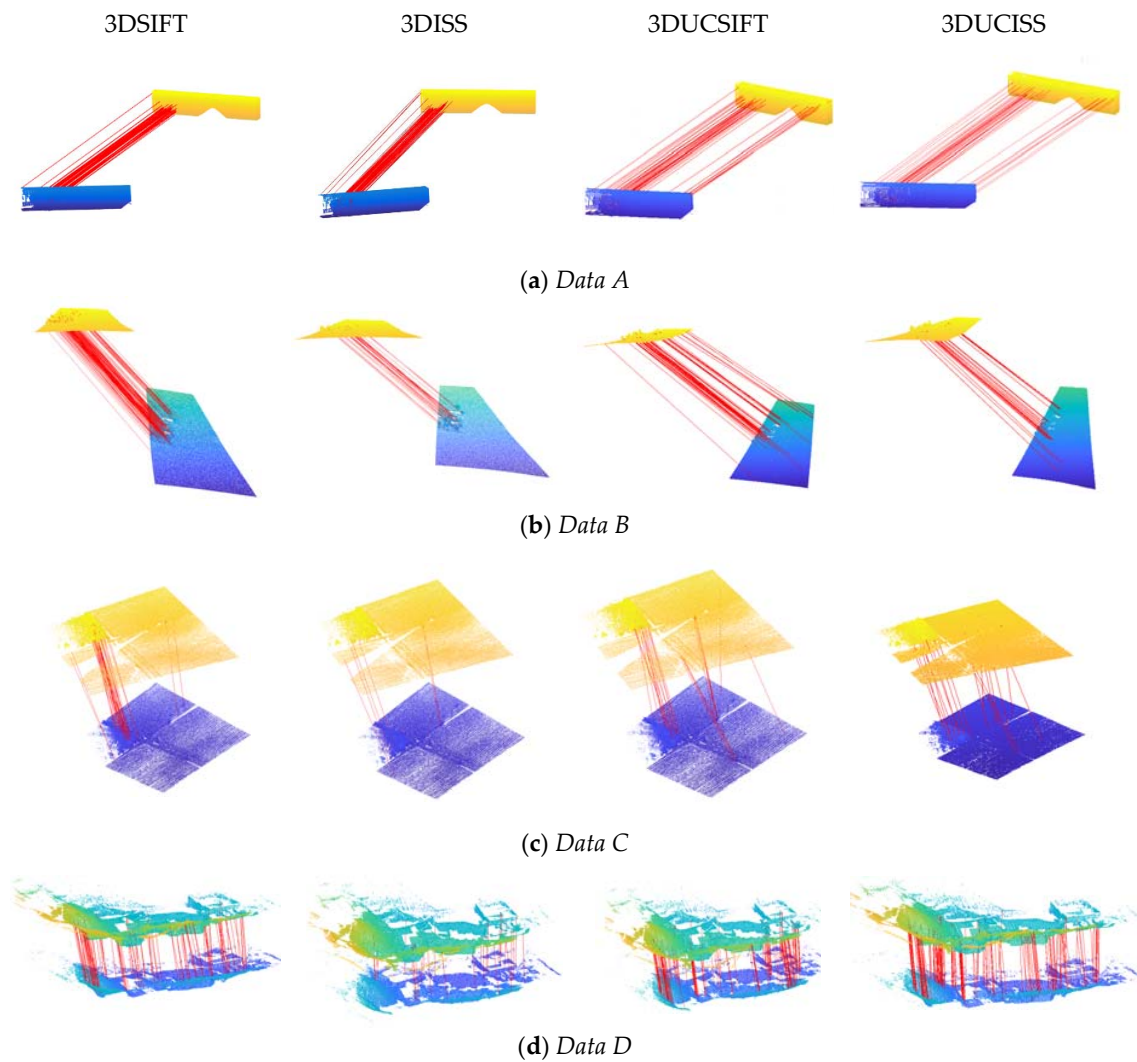


Figure 10. Displaying of correspondence process using 3D SIFT, 3D ISS standard detectors, and the proposed method.

Table 2. Results of comparison of the proposed method and standard methods. mr is the mean of resolution.

		Data A ($mr = 5$ cm)				Data B ($mr = 60$ cm)			
		3DSIFT	3DUCSIFT	3DISS	3DUCISS	3DSIFT	3DUCSIFT	3DISS	3DUCISS
Number of extracted features	Source point cloud	1197	1200	1221	1200	701	950	680	800
	Target point cloud	1188	1200	3744	1200	789	950	773	800
		Data C ($mr = 45$ cm)				Data D ($mr = 15$ cm)			
		3DSIFT	3DUCSIFT	3DISS	3DUCISS	3DSIFT	3DUCSIFT	3DISS	3DUCISS
Number of extracted features	Source point cloud	1391	1500	1057	1500	2452	2500	1700	1700
	Target point cloud	1599	1500	1758	1500	3007	3000	1803	1700

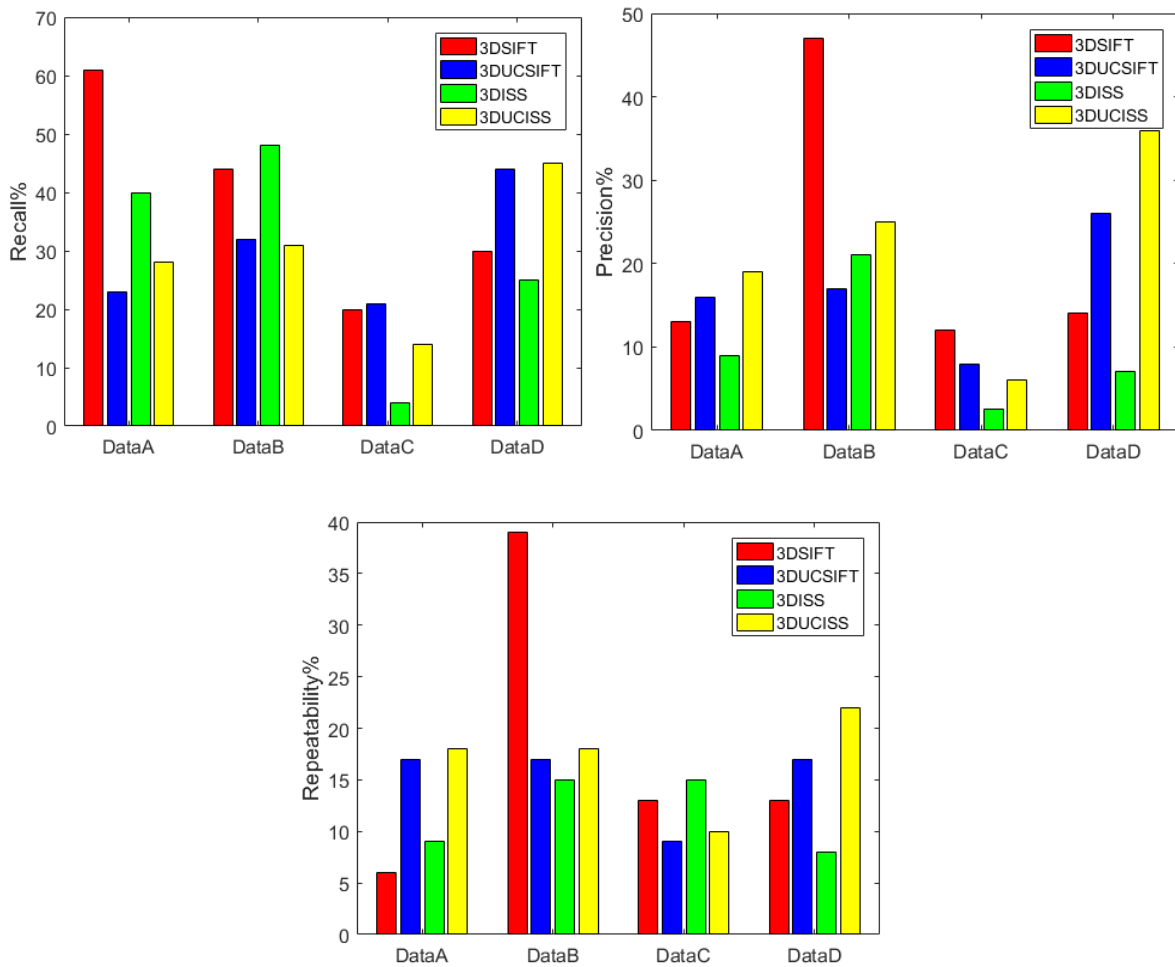


Figure 11. Results of Recall, Precision, and Repeatability criteria.

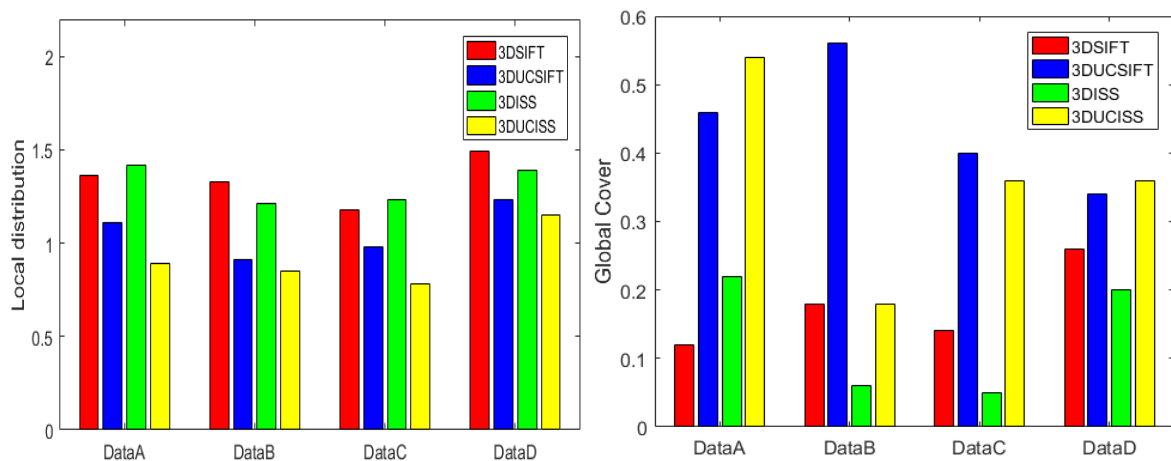


Figure 12. Results of spatial distribution criteria.

According to Figure 11, in some comparisons, we see better performance of the proposed method, and in some cases, standard algorithms are superior in the recall, precision, or repeatability criteria. For example, the 3DSIFT method in data A and B has a higher recall rate, and in data C, this value is very close to the proposed method. This superiority is because of the type of data used in this research. These data consist of a large homogeneous region with little normal angle changes and a cluttered area with severe normal angle changes in a local neighborhood. Due to the nature of the standard keypoint detection

algorithms, the extracted keypoints are often concentrated in a cluttered region of the data, and the keypoint distribution criterion is not considered. As a result, the high Recall obtained by standard detectors algorithms is limited only to correct correspondences in a small area of point clouds. This issue is visible in Figure 10.

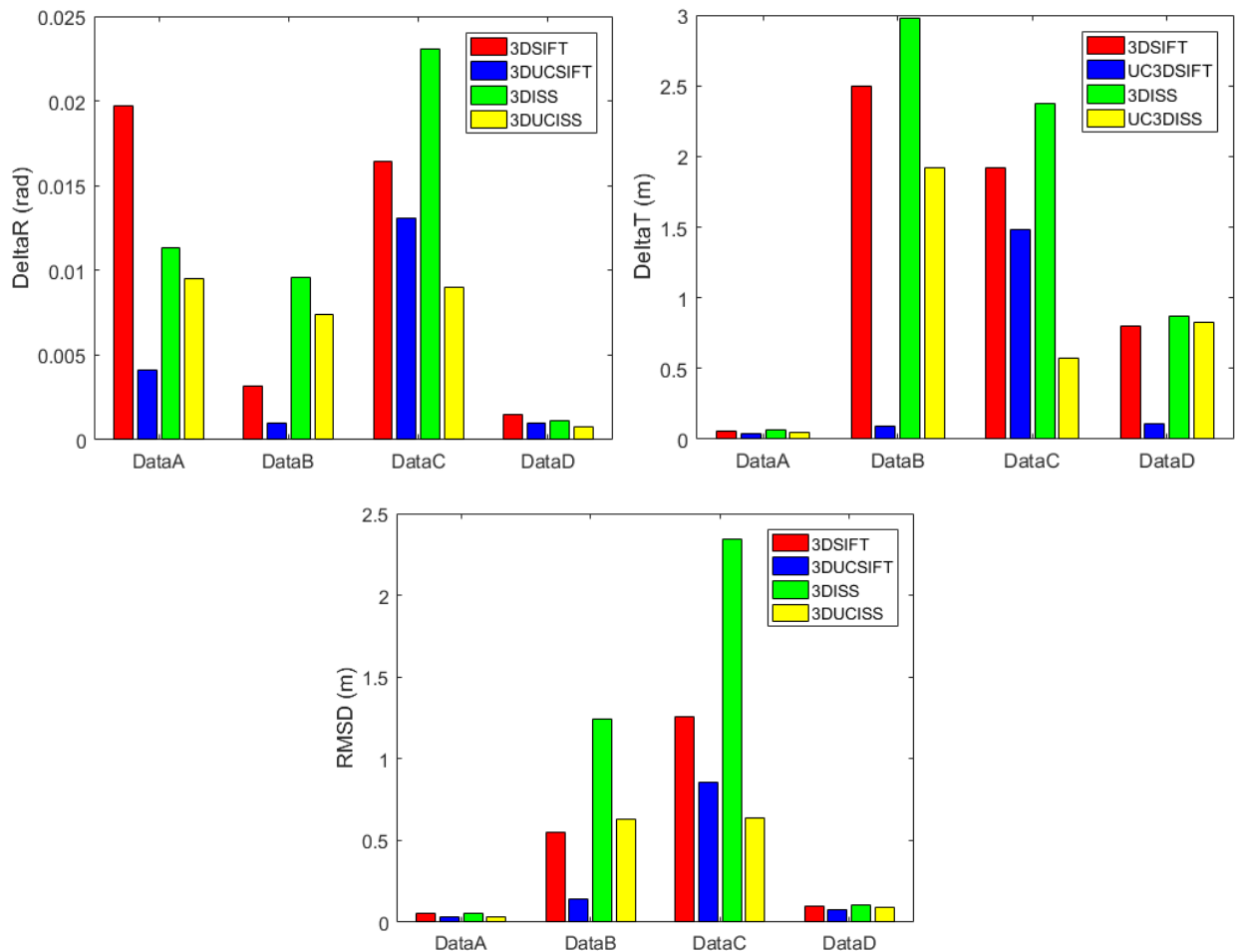


Figure 13. Results of positional accuracy criteria.

According to Figure 12, the proposed method obtains the best performance in the local distribution density and the global coverage index in all data. These results show that the proposed method is able to well consider the distribution of extracted keypoints.

The results of Figure 13 show that the proposed method has a higher performance in all data in terms of the positional accuracy in point cloud registration. The superiority of the proposed method is very noticeable in data A, B, and C, but in data D, this superiority is not significant. In data D, the homogeneous areas are less than in the other data sets used. The results obtained from these data show that the Recall, Precision, and Repeatability criteria in the proposed method are higher than the standard methods. Although the amount of the DeltaT, DeltaR, and RMSD indicate the potential accuracy of the proposed method is more than standard methods, this superiority is not as impressive as for data A, B, and C. The spatial distribution of extracted keypoints by standard detectors is better than in the other data, so the criteria for the positional accuracy in the proposed approach and standard algorithms are almost close. This result testifies to the very high importance of the spatial distribution of the corresponding points in the point cloud. According to the RMSD criterion, the proposed method, in comparison to standard algorithms, has increased the positional accuracy of registration for A, B, C, and D data by 40%, 62%, 52%, and 15%, respectively.

The reason for the superiority of the proposed method over standard algorithms is the extraction of keypoints in a suitable spatial distribution in the point clouds, which significantly increases the registration accuracy in these types of data. On the other hand, another factor influencing the point cloud registration in these kinds of data is the used descriptor. Because many of the extracted points in this type of data are in areas with low normal angle changes, resulting in ambiguity in the matching process. In the proposed method, the most competent keypoints in each cell are detected to make discriminative descriptors and therefore increase the chances of success in the matching step.

5. Conclusions and Suggestions

Three-dimensional keypoint detection in the point clouds with homogeneous structures by standard detectors faces challenges. In this research, challenges such as the controllability of the number of points, quality, and spatial distribution in the 3D keypoint detection process have been studied. The method proposed in this study is tested and analyzed in four different point clouds. The dominant structure of data used was homogeneous, and fewer parts of them were cluttered regions. The core of the proposed approach is the extraction of uniform and competency keypoints to the coarse registration of point clouds with a homogeneous structure. In this research, we tried to identify the keypoints with the highest quality. The higher the quality of the keypoints (the more distinctive they are), the higher the chances of the matching process in the next step. Finally, the registration process, which is the final output, is performed with higher accuracy. For this purpose, a combination of several quality criteria has been used. The reason for combining these criteria is to reach high-quality (distinctive) points from different perspectives. The quality of the keypoints was evaluated using 3D local shape features, 3D local self-similarity, and a histogram of normal orientation. We provided a competency index by a combination of these features. In addition, the spatial distribution of keypoints was made using the Octree structure. In addition to creating suitable spatial distribution for the extracted keypoints, the Octree structure provides more chances to more distinctive points in the cluttered area. In the proposed method, the ability to control the number of extracted keypoints was obtained using Equation (15). The experimental results showed the proper performance of the proposed method in the registration of point clouds with a homogeneous structure. In comparison to the standard algorithms, registration error was diminished by up to 56% by our method. As for future research, the authors suggest the development of the proposed method on other 3D detectors. Our method detects 3D keypoints with appropriate spatial distribution on the whole point cloud. It will be an advantage in dealing with point clouds that have a small overlap. As a result, investigating the performance of the proposed method on data with little overlaps is suggested as future works. Furthermore, developing a stable descriptor for adapting to point clouds with a homogeneous structure can increase the accuracy of the point clouds registration.

Author Contributions: Conceptualization, F.G., H.E., N.P. and A.S.; methodology, F.G., H.E., N.P. and A.S.; validation, F.G. and H.E.; investigation, F.G., H.E. and N.P.; resources, F.G., H.E., N.P. and A.S.; data curation, F.G. and N.P.; writing—original draft preparation, F.G. and H.E.; writing—review and editing, F.G., H.E. and N.P.; visualization, F.G. and H.E.; supervision, H.E., N.P. and A.S.; project administration, F.G.; funding acquisition, N.P.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No data and code available.

Acknowledgments: The authors would like to thank Markus Hollaus for providing ALS data.

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

References

1. Li, W.; Cheng, H.; Zhang, X. Efficient 3D Object Recognition from Cluttered Point Cloud. *Sensors* **2021**, *21*, 5850. [[CrossRef](#)] [[PubMed](#)]
2. Cheng, Q.; Sun, P.; Yang, C.; Yang, Y.; Liu, P.X. A morphing-Based 3D point cloud reconstruction framework for medical image processing. *Comput. Methods Programs Biomed.* **2020**, *193*, 105495. [[CrossRef](#)] [[PubMed](#)]
3. Liang, X.; Hyypä, J.; Kaartinen, H.; Lehtomäki, M.; Pyörälä, J.; Pfeifer, N.; Holopainen, M.; Broly, G.; Francesco, P.; Hackenberg, J.; et al. International benchmarking of terrestrial laser scanning approaches for forest inventories. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 137–179. [[CrossRef](#)]
4. Milenković, M.; Pfeifer, N.; Glira, P. Applying terrestrial laser scanning for soil surface roughness assessment. *Remote Sens.* **2015**, *7*, 2007–2045. [[CrossRef](#)]
5. Theiler, P.W.; Wegner, J.D.; Schindler, K. Keypoint-based 4-points congruent sets—automated marker-less registration of laser scans. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 149–163. [[CrossRef](#)]
6. Besl, P.J.; McKay, N.D. Method for registration of 3-D shapes. In Proceedings of the Sensor Fusion IV: Control Paradigms and Data Structures, Boston, MA, USA, 14–15 November 1991.
7. Glira, P.; Pfeifer, N.; Briese, C.; Ressel, C. A correspondence framework for ALS strip adjustments based on variants of the ICP algorithm. *Photogramm. Fernerkund. Geoinf.* **2015**, *4*, 275–289. [[CrossRef](#)]
8. Yang, J.; Cao, Z.; Zhang, Q. A fast and robust local descriptor for 3D point cloud registration. *Inf. Sci.* **2016**, *346*, 163–179. [[CrossRef](#)]
9. Salti, S.; Tombari, F.; di Stefano, L. A performance evaluation of 3d keypoint detectors. In Proceedings of the 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, Hangzhou, China, 16–19 May 2011.
10. Shah, S.A.A.; Bennamoun, M.; Boussaid, F. Keypoints-based surface representation for 3D modeling and 3D object recognition. *Pattern Recognit.* **2017**, *64*, 29–38. [[CrossRef](#)]
11. Tang, J.; Ericson, L.; Folkesson, J.; Jensfelt, P. GCNv2: Efficient correspondence prediction for real-time SLAM. *IEEE Robot. Autom. Lett.* **2019**, *4*, 3505–3512. [[CrossRef](#)]
12. Nie, W.; Xiang, S.; Liu, A. Multi-scale CNNs for 3D model retrieval. *Multimed. Tools Appl.* **2018**, *77*, 22953–22963. [[CrossRef](#)]
13. Wang, Y.; Yang, B.; Chen, Y.; Liang, F.; Dong, Z. JoKNet: A joint keypoint detection and description network for large-scale outdoor TLS point clouds registration. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *104*, 102534.
14. Yoshiki, T.; Kanji, T.; Naiming, Y. Scalable change detection from 3d point cloud maps: Invariant map coordinate for joint viewpoint-change localization. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018.
15. Tombari, F.; Salti, S.; di Stefano, L. Performance evaluation of 3D keypoint detectors. *Int. J. Comput. Vis.* **2013**, *102*, 198–220. [[CrossRef](#)]
16. Chen, H.; Bhanu, B. 3D free-form object recognition in range images using local surface patches. *Pattern Recognit. Lett.* **2007**, *28*, 1252–1262. [[CrossRef](#)]
17. Zhong, Y. Intrinsic shape signatures: A shape descriptor for 3d object recognition. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, Kyoto, Japan, 27 September–4 October 2009.
18. Mian, A.; Bennamoun, M.; Owens, R. On the repeatability and quality of keypoints for local feature-based 3d object retrieval from cluttered scenes. *Int. J. Comput. Vis.* **2010**, *89*, 348–361. [[CrossRef](#)]
19. Fiolka, T.; Stückler, J.; Klein, D.A.; Schulz, D.; Behnke, S. Sure: Surface entropy for distinctive 3d features. In *International Conference on Spatial Cognition*; Springer: Berlin/Heidelberg, Germany, 2012.
20. Sipiran, I.; Bustos, B. Harris 3D: A robust extension of the Harris operator for interest point detection on 3D meshes. *Vis. Comput.* **2011**, *27*, 963–976. [[CrossRef](#)]
21. Castellani, U.; Cristani, M.; Fantoni, S.; Murino, V. Sparse points matching by combining 3D mesh saliency with statistical descriptors. In *Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2008.
22. Zaharescu, A.; Boyer, E.; Varanasi, K.; Horaud, R. Surface feature detection and description with applications to mesh matching. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
23. Donoser, M.; Bischof, H. 3d segmentation by maximally stable volumes (msvs). In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, 20–24 August 2006.
24. Rusu, R.B.; Cousins, S. 3d is here: Point cloud library (pcl). In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011.
25. Persad, R.A.; Armenakis, C. Automatic co-registration of 3D multi-sensor point clouds. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 162–186. [[CrossRef](#)]
26. Weinmann, M.; Weinmann, M.; Hinz, S.; Jutzi, B. Fast and automatic image-based registration of TLS data. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, S62–S70. [[CrossRef](#)]
27. Petricek, T.; Svoboda, T. Point cloud registration from local feature correspondences—Evaluation on challenging datasets. *PLoS ONE* **2017**, *12*, e0187943. [[CrossRef](#)]
28. Bueno, M.; González-Jorge, H.; Martínez-Sánchez, J.; Lorenzo, H. Automatic point cloud coarse registration using geometric keypoint descriptors for indoor scenes. *Autom. Constr.* **2017**, *81*, 134–148. [[CrossRef](#)]

29. Zhu, J.; Xu, Y.; Hoegner, L.; Stilla, U. Direct co-registration of TIR images and MLS point clouds by corresponding keypoints. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *4*, 235–242. [[CrossRef](#)]
30. Tonioni, A.; Salti, S.; Tombari, F.; Spezialetti, R.; Stefano, L.D. Learning to detect good 3D keypoints. *Int. J. Comput. Vis.* **2018**, *126*, 1–20. [[CrossRef](#)]
31. Teran, L.; Mordohai, P. 3D interest point detection via discriminative learning. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2014.
32. Alcantarilla, P.F.; Bartoli, A.; Davison, A.J. KAZE features. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2012.
33. Salti, S.; Tombari, F.; Spezialetti, R.; Di Stefano, L. Learning a descriptor-specific 3D keypoint detector. In Proceedings of the IEEE International Conference on Computer Vision, Washington, DC, USA, 7–13 December 2015.
34. Godil, A.; Wagan, A.I. Salient local 3D features for 3D shape retrieval. In *Three-Dimensional Imaging, Interaction, and Measurement*; SPIE: Bellingham, MA, USA, 2011.
35. Hänsch, R.; Weber, T.; Hellwich, O. Comparison of 3D interest point detectors and descriptors for point cloud fusion. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *2*, 57. [[CrossRef](#)]
36. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
37. Salti, S.; Tombari, F.; di Stefano, L. SHOT: Unique signatures of histograms for surface and texture description. *Comput. Vis. Image Underst.* **2014**, *125*, 251–264.
38. Hana, X.F.; Jin, J.S.; Xie, J.; Wang, M.J.; Jiang, W. A comprehensive review of 3D point cloud descriptors. *arXiv* **2018**, arXiv:1802.02297.
39. Weinmann, M. *Reconstruction and Analysis of 3D Scenes*; Springer: Berlin/Heidelberg, Germany, 2016.
40. Huang, J.; You, S. Point cloud matching based on 3D self-similarity. In Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Providence, RI, USA, 16–21 June 2012.
41. Prakhya, S.M.; Liu, B.; Lin, W. Detecting keypoint sets on 3D point clouds via Histogram of Normal Orientations. *Pattern Recognit. Lett.* **2016**, *83*, 42–48. [[CrossRef](#)]
42. Samet, H. An overview of quadtrees, octrees, and related hierarchical data structures. *Theor. Found. Comput. Graph. CAD* **1988**, *40*, 51–68.
43. Vo, A.V.; Truong-Hong, L.; Laefer, D.F.; Bertolotto, M. Octree-based region growing for point cloud segmentation. *ISPRS J. Photogramm. Remote Sens.* **2015**, *104*, 88–100. [[CrossRef](#)]
44. Lu, B.; Wang, Q.; Li, A.N. Massive Point Cloud Space Management Method Based on Octree-Like Encoding. *Arab. J. Sci. Eng.* **2019**, *44*, 9397–9411. [[CrossRef](#)]
45. Park, S.; Ju, S.; Yoon, S.; Nguyen, M.H.; Heo, J. An efficient data structure approach for BIM-to-point-cloud change detection using modifiable nested octree. *Autom. Constr.* **2021**, *132*, 103922. [[CrossRef](#)]
46. Eggert, D.; Dalyot, S. Octree-based simd strategy for icp registration and alignment of 3d point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *1*, 105–110. [[CrossRef](#)]
47. Sedaghat, A.; Mohammadi, N. Uniform competency-based local feature extraction for remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2018**, *135*, 142–157. [[CrossRef](#)]
48. Sedaghat, A.; Mokhtarzade, M.; Ebadi, H. Uniform robust scale-invariant feature matching for optical remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4516–4527. [[CrossRef](#)]
49. Mousavi, V.; Varshosaz, M.; Remondino, F. Using Information Content to Select Keypoints for UAV Image Matching. *Remote Sens.* **2021**, *13*, 1302. [[CrossRef](#)]
50. Berretti, S.; Werghi, N.; Del Bimbo, A.; Pala, P. Selecting stable keypoints and local descriptors for person identification using 3D face scans. *Vis. Comput.* **2014**, *30*, 1275–1292. [[CrossRef](#)]
51. Ghorbani, F.; Ebadi, H.; Sedaghat, A.; Pfeifer, N. A Novel 3-D Local DAISY-Style Descriptor to Reduce the Effect of Point Displacement Error in Point Cloud Registration. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 2254–2273. [[CrossRef](#)]
52. Stancelova, P.; Sikudova, E.; Cernekova, Z. 3D Feature detector-descriptor pair evaluation on point clouds. In Proceedings of the 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, The Netherlands, 18–21 January 2021.
53. Dong, Z.; Yang, B.; Liang, F.; Huang, R.; Scherer, S. Hierarchical registration of unordered TLS point clouds based on binary shape context descriptor. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 61–79. [[CrossRef](#)]