Fast range-independent spherical subsampling of 3D laser scanner points and data reduction performance evaluation for scene registration

Anthony Mandow, Jorge L. Martínez, Antonio J. Reina and Jesús Morales

Abstract: Three-dimensional laser range-finders are increasingly being incorporated into applications, such as mobile robotics, that require real-time registration of scene data. However, the computational costs of adaptive range-dependent data selection and point cloud matching grow significantly with the number of points. Therefore, fast range-independent subsampling by uniform or random data reduction is usually performed at a preprocessing step. The paper proposes a new range-independent subsampling algorithm that is more effective for the widely used spherical scanning mechanism. As this type of device measures the ranges by composition of two rotations, it samples certain directions with a higher density, which can distort the registration optimization process. The proposed solution uses sensor characteristics to equalize the measure-direction density of the reduced point cloud. The paper also addresses performance assessment of subsampling methods by contributing three benchmark criteria that do not rely on a particular registration technique: one considers the ground truth transformation between two scans, and the other two are based on the analysis of a single scan. The advantages of spherical subsampling are analyzed through a comparison of range-independent methods and a simple range-dependent one with real scans from three representative scenes (urban, natural, and indoors).

Keywords: 3D measurement system; Laser ranging; Point subsampling; Scene registration; Mobile robotics; Point matching

How to Cite:

Mandow, A., Martínez, J.L., Reina, A.J., Morales, J.

© 2010 Copyright by the authors.
Fast Range-Independent Spherical Subsampling of 3D Laser Scanner Points and Data Reduction Performance Evaluation for Scene Registration.

Anthony Mandow*, Jorge L. Martínez, Antonio J. Reina and Jesús Morales

Departamento de Ingeniería de Sistemas y Automática, E.T.S. Ingenieros Industriales
Universidad de Málaga, 29071 Málaga, Spain

Abstract

Three-dimensional laser range-finders are increasingly being incorporated into applications, such as mobile robotics, that require real-time registration of scene data. However, the computational costs of adaptive range-dependent data selection and point cloud matching grow significantly with the number of points. Therefore, fast range-independent subsampling by uniform or random data reduction is usually performed at a preprocessing step. The paper proposes a new range-independent subsampling algorithm that is more effective for the widely used spherical scanning mechanism. As this type of device measures the ranges by composition of two rotations, it samples certain directions with a higher density, which can distort the registration optimization process. The proposed solution uses sensor characteristics to equalize the measure-direction density of the reduced point cloud. The paper also addresses performance assessment of subsampling methods by contributing three benchmark criteria that do not rely on a particular registration technique: one considers the ground truth transformation between two scans, and the other two are based on the analysis of a single scan. The advantages of spherical subsampling are analyzed through a comparison of range-independent methods and a simple range-dependent one with real scans from three representative scenes (urban, natural, and indoors).

Key words: 3D measurement system, Laser ranging, Point subsampling, Scene registration, Mobile robotics, Point matching, Performance Evaluation

1. Introduction

Three-dimensional (3D) laser range-finders are increasingly being incorporated into applications that require fast and efficient scene registration, such as virtual modeling of urban [1] and disaster areas [2], surveynce for change detection [3], and autonomous mobile robot navigation [4] [5]. These sensors scan the scene to obtain a collection of ranges that can be represented as a matrix of distances to scene points (i.e., a range image).

Two scanning procedures are usually employed for 3D data acquisition. First, a 2D laser device can be displaced either by an automated system, for object modelling [6][7], or by a vehicle, for urban modelling [1][8]. Second, ranges can be obtained from a fixed pose by composition of two rotations of the laser beam, which renders spherical coordinates [9]. With spherical scanners, commonly found in mobile robotics, certain scene regions are scanned with a higher density, which can distort the registration optimization process.

*This work was partially supported by the Spanish CICYT project DPI 2008-00533.

**THIS IS A PREPRINT. The published version of the article is available in: The Pattern Recognition Letters 31 (2010) 1239-1250. doi:10.1016/j.patrec.2010.03.008

*Corresponding author Email: amandow@uma.es, Tel: (+34) 951 952 250.

Preprint submitted to Pattern Recognition Letters January 27, 2011
Registration is an optimization problem that searches for the spatial transformation of a source range image that produces maximum overlap with a destination scan. Real world scenes can be populated by uneven surfaces and multiform objects, such as stairs, trees, or cars. Even though registration can rely on geometric features (e.g., planes) for urban [10] and indoor environments [8], the use of raw range data (i.e., point clouds) provides a more general approach for less structured (e.g., natural) environments.

Registration of a pair of point clouds requires establishing point-wise spatial correspondences with 6 degrees of freedom [11], so the computational cost of the optimization problem grows with the number of points [12]. Regardless of the efforts devoted to increase both hardware performance [13] and the efficiency of registration algorithms [14], fast and effective subsampling of scan data remains a key preprocessing step for point cloud registration [6]. Nevertheless, even if 3D scan data reduction is often taken for granted, subsampling methods have not received enough attention in literature, with the exception of the brief review found in [14].

Laser data subsampling methods can be broadly classified into two types:

- **Range-dependent**, where each scan is processed to find feature points with special characteristics. These can be representative points from sequences with similar ranges in the same 2D slice [15] or points from salient geometrical regions [4]. Other approaches choose points from a spatial subdivision of the point cloud, such as mesh vertices [16] or octree cube centers [17]. Furthermore, selection can be made from attributes computed for every single point, either as geometric descriptors, like surface curvature [18] and normals [14], or as more complex features [19][20]. In addition range-dependent strategies can be adopted to attain a uniform point density on some scene surfaces [2]. In general, range-driven techniques can select a representative set of points, but they require the computational load of processing information content for every scan and can be affected by sensor range errors.

- **Range-independent**. Subsampling can be implemented as a pre-computed binary mask applied to the range image. Computationally simple systematic methods can be used either alone or as a first step in combination with range-dependent methods. Typically, range-independent subsampling has relied on random [21] or uniform [12] [22] selection strategies, which do not consider the scanning principle. Spherical subsampling was introduced in [23] to equalize the measure-direction density according to the sensor resolution of spherical scanners. This equalization has an implicit data reduction rate.

This paper analyzes fast data reduction methods for 3D range scans. In particular, the major motivations of this work are the following:

- To extend the original concept of spherical point subsampling [23] to allow for any further data reduction by considering spherical scanner characteristics.

- To define performance criteria for subsampling methods that can be independent of a particular point matching technique.

- To experimentally compare simple subsampling methods in environments and conditions found in mobile robotics.

The paper is organized as follows. Next section enunciates the spherical subsampling method. Section 3 proposes performance criteria for subsampling methods. Section 4 illustrates the application of different subsampling methods to 3D scans from three case study scenes. Section 5 applies performance criteria to analyze and compare subsampling methods with the case study scans. Finally, conclusions and references complete the paper.

2. Spherical Subsampling

2.1. 3D Scene Range Data

Three types of 3D laser devices are commonly employed to obtain range data:
Commercial short range 3D devices with reduced field of view, usually applied for reverse engineering or for generating computer graphics from small objects [6][7][16].

High-end long range 3D devices with high resolution employed for realistic visualization and documentation of sites of interest [1][24].

Standard 2D devices customized with an extra degree freedom [9]. This solution is commonly adopted in mobile robotics [4][5][15] and also in the experimental setup that illustrates this work.

For scene perception, sensors usually belong to one of the last two categories, where a fast 2D scan can be combined in several ways with an additional rotation of the whole opto-mechanical head [9][24]. Thus, each point is acquired in spherical coordinates given by the measured range and two rotations. This is illustrated in Fig. 1(a) for a standard 180° 2D scanner, whose optical head rotates around its local Z axis to measure polar distances on its XY plane, with the Y axis given as the bisector of the 2D scan angle. In the figure, the additional degree of freedom consists on pivoting the 2D sensor around its X axis. The 3D sensor frame is defined so that it coincides with that of the 2D range-finder for the middle position of the new rotation. The paper focuses only on this case, but the proposed subsampling method can be adapted easily to other combinations, such as a rotation about the Y axis.

The field of view of the 3D sensor is defined by the scope of both scan angles: Φ for the 2D device and Ψ for the additional rotation, as shown in Fig. 1(a). Angular resolution is given by Δφ and Δψ for the rotations around Z and X, respectively. Thus, a 3D scan consists of \( N_\psi \) 2D slices each of which is composed of \( N_\phi \) laser readings:

\[
N_\psi = 1 + \frac{\Psi}{\Delta \psi}; \quad N_\phi = 1 + \frac{\Phi}{\Delta \phi}. \quad (1)
\]

Dots in Fig. 1(a) represent measurement directions (or lines of view) with respect to the origin of the scanner coordinate system. Note that even though Δφ and Δψ have constant values, the density of measurement directions is not homogeneous. This is because the arc length swept by Ψ is shorter when the scan direction is closer to the additional rotation axis [9]. Thus, for a sensor with \( \Psi = 180^\circ \), the first direction of all 2D scans coincides with the X axis, and the same is true for the last direction. This means that all but one element of the first and the last columns of matrix \( R \) would be redundant for a static scene. Scan direction densities for each column get sparser when the direction approaches the YZ plane (i.e., in the forward direction of the sensor, where column resolution actually becomes \( \Delta \psi \)).
2.2. Subsampling Procedure

The goal of the spherical subsampling method is to reduce scan data by pursuing an equalization of measure-direction densities with a reference angular resolution $\Delta \theta_s$. This is illustrated in Fig. 1(b), where 'o's denote range directions selected from the raw scan in Fig. 1(a).

The input parameter for the subsampling procedure $\Delta \theta_s$ can be specified by considering the original sensor sampling resolutions as follows:

$$\Delta \theta_s = \max(\Delta \psi, \Delta \phi) p,$$

where $0 < p \leq 1$ is an equalization factor. In this way, by setting $p = 1$, the measure-direction density is equalized with the coarsest resolution actually provided by the sensor.

As the measure-direction density in a 2D scan is uniform, the same number of columns $n_c$ is kept for every row in matrix $R$:

$$n_c = \text{round}\left(1 + \frac{\Phi}{\Delta \theta_s}\right).$$

Then, the set of selected column indexes $J$ is the same for all rows. This set is obtained uniformly as follows:

$$J = \left\{1 + \text{round}\left(\frac{N \phi - 1}{s - 1} (k - 1)\right)\right\}_{k=1..n_c}.$$

The measure-direction density of the columns in $R$ varies according to the corresponding swept arc length, which is maximum for the central column (see Fig. 1(a)). A normalized arc length $0 \leq a_j \leq 1$ can be defined for the $j$-th column as:

$$a_j = \sin\left(\pi - \frac{\Phi}{2} + (j - 1) \Delta \phi\right),$$

where $\Phi \leq \pi$. Thus, the number $n_r(j)$ of row points chosen for the $j$-th column is a function of $a_j$:

$$n_r(j) = \begin{cases} \text{round}\left(1 + \frac{\Phi}{N \phi} - 1\right) a_j & \text{if } j \in J, \\ 0 & \text{otherwise.} \end{cases}$$

Then, the set $I_j$ of selected row indices for the $j$-th column is:

$$I_j = \left\{1 + \text{round}\left(\frac{N \phi - 1}{n_r(j)} (k - 1)\right)\right\}_{k=1..n_r(j)}.$$

A binary mask matrix $M$ is defined by setting to one the elements indexed by (4) and (7). This matrix, which represents the points selected from $R$, has to be computed just once for a particular scanner configuration. Then, element-by-element multiplication of mask $M$ with scan $R$ will systematically produce a subsampled matrix $R_s$, where null elements will not be used for registration.

The total subsampling proportion $p_t$ indicates the rate of selected points with respect to the original 3D scan. This rate is given by:

$$p_t = \frac{\sum_{j=1}^{N} n_r(j)}{N}. $$

The value of $p_t$ for $p = 1$ in (2) is the characteristic subsampling proportion of the spherical method for a given sensor. Any further data reduction can be achieved with smaller values of $p$. 

4
3. Subsampling Performance Criteria

The assessment of a subsampling method for registration is a nontrivial problem. Usually, performance has been assessed in terms of matching accuracy [25] and geometric convergence [26] by considering the registration process as a whole. This incorporates errors accumulated by the different stages of a particular registration procedure, so subsampling is not considered per se.

This section proposes subsampling performance criteria that do not rely on a given point registration technique. The first two criteria are based on the analysis of a single scan, so they do not require ground truth knowledge. The third one considers the ground truth transformation between two scans.

3.1. Spatial Distribution Index

This index is based on the 3D binary grid concept, which is a collection of constant size cube-shaped binary cells. The boolean value of a cell is set to one if at least one scan point, in Cartesian coordinates, falls within its limits. The set of one-valued cells for a subsampled scan is a subset of the set of one-valued cells for the complete scan. Then, given a grid resolution, the ratio \( P_{sd} \) of one-valued cells between the subsampled and the original scans represents a measurement of the preservation of 3D spatial distribution.

Local Cartesian coordinates for a range matrix element \( r_{i,j} \) can be obtained as:

\[
\begin{bmatrix}
x_{i,j} \\
y_{i,j} \\
z_{i,j}
\end{bmatrix} = r_{i,j} \begin{bmatrix}
\cos(i \Delta \phi) \\
\cos(j \Delta \psi) \sin(i \Delta \phi) \\
\sin(j \Delta \psi) \sin(i \Delta \phi)
\end{bmatrix}
\]

(9)

The 3D grid can be represented as a binary matrix \( G \) of dimensions \( g_x \times g_y \times g_z \). For a cell edge length \( L \), the dimensions of the matrix that envelopes all scan readings are:

\[
g_x = \text{round}((x_{\text{max}} - x_{\text{min}})/L),
\]

\[
g_y = \text{round}((y_{\text{max}} - y_{\text{min}})/L),
\]

\[
g_z = \text{round}((z_{\text{max}} - z_{\text{min}})/L),
\]

where \( x_{\text{max}}, y_{\text{max}}, z_{\text{max}} \), and \( x_{\text{min}}, y_{\text{min}}, z_{\text{min}} \) are the highest and lowest bounds, respectively, for all scanned points in Cartesian coordinates, as obtained with (9).

Then, the corresponding element in \( G \) is set to one by scan point \( [x_{i,j}, y_{i,j}, z_{i,j}]^T \), as follows:

\[
G \left( \text{round} \left( \frac{y_{i,j} - y_{\text{min}}}{L} \right), \text{round} \left( \frac{y_{i,j} - y_{\text{min}}}{L} \right), \text{round} \left( \frac{y_{i,j} - y_{\text{min}}}{L} \right) \right) = 1.
\]

(11)

Let \( G_R \) and \( G_{R_s} \) be the resulting \( G \) matrices for all the points in the range matrix \( R \) and for the subsampled points in range matrix \( R_s \), respectively. Note that both matrices have the same dimensions and that all non-zero elements in \( G_R \) are also non-zero elements in \( G_{R_s} \). Then, the spatial distribution index is given by:

\[
P_{sd}(L) = \frac{\text{nnz}(G_{R_s})}{\text{nnz}(G_R)}
\]

(12)

where \( \text{nnz()} \) is the function that returns the total number of non-zero elements. This is a normalized index, where values closer to 1 would indicate a better preservation of 3D spatial distribution after scan data reduction, by considering a spatial resolution \( L \). The values of this index for a range of spatial resolutions can be considered to assess subsampling performance.

3.2. Matchability After Transformation Index

When scans are taken from a moving platform (e.g., a mobile robot), not all points are matchable due to occlusions, point density changes, or range image limits. Thus, a performance index for scan registration suitability can be defined as the proportion of subsampled points that remain matchable after a given spatial transformation of the sensor.
These can be found by projecting the global Cartesian coordinates of subsampled points, as obtained from (9), onto a spatial transformation of the range image. If the sensor is subject to a translation \([x_0, y_0, z_0]^T\) and a Roll-Pitch-Yaw rotation \([\alpha, \beta, \gamma]^T\), then the local coordinates of the source scan points will be changed according to:

\[
\begin{bmatrix}
\hat{x}_{i,j} \\
\hat{y}_{i,j} \\
\hat{z}_{i,j}
\end{bmatrix} = \begin{bmatrix}
x_0 \\
y_0 \\
z_0
\end{bmatrix} + \begin{bmatrix}
C_\alpha C_\beta & -S_\alpha C_\gamma + C_\alpha S_\beta S_\gamma & S_\alpha S_\gamma + C_\alpha S_\beta C_\gamma \\
S_\alpha C_\beta & C_\alpha C_\gamma + S_\alpha S_\beta S_\gamma & -C_\alpha S_\gamma + S_\alpha S_\beta C_\gamma \\
-S_\beta & C_\beta S_\gamma & C_\beta C_\gamma
\end{bmatrix} \begin{bmatrix}
x_{i,j} \\
y_{i,j} \\
z_{i,j}
\end{bmatrix},
\]  

(13)

where \(C\) and \(S\) denote the cosine and sine functions, respectively.

To project the non-zero points of \(R_s\) onto the new range matrix \(\hat{R}\), the matrix indices \((\hat{i}, \hat{j})\) are computed as follows:

\[
\hat{i} = \text{round}\left(\arctan\left(\frac{\hat{z}_{i,j}}{\hat{x}_{i,j}}\right) \cdot \frac{1}{\Delta\psi}\right) - \Psi_0,
\]

\[
\hat{j} = \text{round}\left(\arctan\left(\frac{\sqrt{\hat{y}_{i,j}^2 + \hat{z}_{i,j}^2}}{\hat{x}_{i,j}}\right) \cdot \frac{1}{\Delta\phi}\right) - \Phi_0,
\]

(14)

where \(\Psi_0\) and \(\Phi_0\) are the angle offsets that correspond to \(i = 0\) and \(j = 0\), respectively.

If \(0 \leq \hat{i} \leq N_\psi\) and \(0 \leq \hat{j} \leq N_\phi\), the new coordinates are within the limits of the transformed range image, and the corresponding range can be computed as:

\[
\hat{r}_{i,j} = \sqrt{\hat{x}_{i,j}^2 + \hat{y}_{i,j}^2 + \hat{z}_{i,j}^2}.
\]

(15)

Note that the limits of the range image are not the only cause for matchability losses. Different ranges in the source scan can be projected onto the same element of the destination range image, as computed by (14), because the number of points subtended by a certain object can be reduced if the sensor gets farther from it. Moreover, if a point becomes occluded by other after the transformation, they will both project onto the same matrix coordinates. To account for these considerations, the projected matrix \(\hat{R}\), which is initialized with zeroes, is constructed by assigning the smallest of the range values obtained for every \((\hat{i}, \hat{j})\) element.

By using (9) and (13) to (15) for a spatial transformation, a normalized index \(P_{\text{mat}}\) can be defined as follows:

\[
P_{\text{mat}} = \frac{\text{nnz}(\hat{R})}{\text{nnz}(R_s)}.
\]

(16)

Then, subsampling performance can be evaluated by computing this index for a range of values in every single translation/rotation transformation.

3.3. Ground Truth Based Performance

A common approach to registration is optimization of the degree of overlap between two scans. This can be computed as the number point matches between a destination and a source scan. Particularly, two threshold-based approaches are usually employed in registration techniques [14]:

- **Closest points.** Points in the source scan are transformed into Cartesian coordinates of the destination scan frame, using (13). They are matched if the distance to their closest destination scan point is below a threshold [12].

- **Reverse calibration.** Source points are matched based on their projection onto the destination range matrix (14) when the difference between the source range (15) and the corresponding destination range is under a given threshold [22].

Thus, if the ground truth transformation is known for a pair of scans, subsampling performance can be assessed as the rate between the number of point matches and the number of subsampled points. This way, two performance indices, \(P_{\text{cp}}(D)\) and \(P_{\text{rc}}(D)\), can be computed for closest points and reverse calibration, respectively, where \(D\) represents a given threshold.
4. Experimental Setup

This section offers the experimental setup that will be employed for subsampling performance evaluation in section 5. It firstly describes the scanner device used in the experiments. Also, three different case study environments are presented. Then, subsampled masks for the sensor are applied to 3D views of these environments. Finally, a simple range-dependent method is described.

4.1. DGPS/IMU-aided 3D Scanner System

A customized DGPS/IMU-aided 3D scanning system (see Fig. 2) has been employed in the experiments. The combination of DGPS, IMU and laser rangefinders is commonly used for terrestrial surveys [27] and mobile robotics [28]. A webcam has been added for documentation purposes. The Inertial Measurement Unit (IMU) combines accelerometers, gyroscopes and magnetometers. The static accuracy of the unit is ±0.75° for roll and pitch angles, and ±2° for yaw.

The DGPS is a dual-frequency Real-Time Kinematic receiver that accepts differential corrections from its local base station via a radio link. It provides horizontal and vertical positioning errors around 1.5 cm and 2 cm, respectively, with good sky visibility.

The 3D scanner device has been constructed by adding an extra degree of freedom to a commercial 2D time-of-flight rangefinder (see Fig. 2). The 2D rangefinder provides up to 81 m range, with ±4 cm range error for ranges of 20 m. It offers a field of view Φ = 180° with horizontal angular resolution Δφ = 1°.

The vertical angular resolution of a 3D scan is Δψ = 1°, with Ψ = 60° of field of view. Thus, the total number of ranges is N = 65341 (361 × 181). The maximum scan volume is contained by a rectangular prism defined by [x_{min} = -81 m, x_{max} = 81 m, y_{min} = 0 m, y_{max} = 81 m, z_{min} = -40.5 m, z_{max} = 40.5 m] on the local sensor axes.
4.2. Case Study Scenes

The results presented hereunder correspond to two outdoor sites (urban and natural) and an indoor scene, whose dimensions exceed the sensor range. These sites offer representative case studies of environment types.

Panoramic photographs of these environments are presented in Fig. 3. Next to each photograph, the scenes are depicted in range matrix \( R \) format as recorded by the laser scanner. 3D data are represented in gray-scale, with closer ranges having a darker gray level. A logarithmic scale has been used so that range details in the foreground can be distinguished in the figure. For illustration purposes, the pixel aspect ratio of all the range images has been set to 3:2 on account of sensor resolutions \( \Delta \phi \) and \( \Delta \psi \).

White points in the range images correspond to the maximum range of the sensor. This indicates either that objects are beyond the sensor limits (e.g., the sky in outdoor scenes) or that the range is erroneous (as in certain surfaces in the urban and indoor environments). Therefore, these limit ranges cannot be employed for assessing subsampling performance.

4.3. Subsampling Masks

Range-independent binary masks have been obtained for the uniform, random, and spherical subsampling methods. The uniform method has been implemented by evenly selecting one out of every \( 1/p_t \) points of the range matrix \( R \). As for the random mask matrix, \( p_tN \) different points of \( R \) have been randomly selected.

Fig. 4 shows the mask matrix \( M \) obtained for different subsampling approaches, where discarded points (i.e., zeroes) are shown in white. The spherical characteristic subsampling proportion (with \( p = 1 \)) for the presented sensor is \( p_t = 0.426 \), as computed from (8). This same subsampling rate has been used for the masks in Fig. 4(a)-(c). The application of these masks to a range image is illustrated in Fig. 5(a)-(c) for the 3D scan of Fig. 3(b). Note that uniform and random masks are independent of the sensor characteristics.

In the case of spherical subsampling (see Fig. 4(c)), point reduction is greater in the side regions, where scan density is higher. In this sensor \( \Delta \psi < \Delta \phi \), so \( \Delta \theta_s \) is obtained in (2) as the horizontal value \( \Delta \theta_s = \Delta \phi = 0.5^\circ \). A spherical mask with an equalization factor \( p = 0.5 \), which yields a total subsampling rate of \( p_t = 0.108 \), is presented in Fig. 4(d). This mask achieves equalization for a coarser resolution of \( \Delta \theta_s = \Delta \phi/0.5 = 1^\circ \).
Figure 4: Binary subsampling masks: (a) Uniform ($p_t = 0.426$). (b) Random ($p_t = 0.426$). (c) Spherical ($p_t = 0.426$). (d) Spherical ($p_t = 0.108$).

Figure 5: Subsampling of the flat roof range image: (a) Uniform ($p_t = 0.426$). (b) Random ($p_t = 0.426$). (c) Spherical ($p_t = 0.426$). (d) Range-dependent ($p_t = 0.425$).
4.4. Range-dependent subsampling

A simple range-dependent subsampling method [15] that can be applied to all kinds of scenes has been implemented for comparison purposes. It selects one point from consecutive readings in the same 2D slice if their ranges are similar. Two parameters can be adjusted: the maximum number of points that can be joined into one, and the threshold that establishes range similarity. These parameters need to be tuned for a given range image so that a desired data reduction is achieved. Thus, to obtain a subsampling proportion similar to that of range-independent masks, point selection for the roof scene shown in Fig. 5(d) has been obtained by considering up to 3 points within a range threshold of 0.1 m. This pattern is recognizable in the rows of the range image.

5. Subsampling Performance Evaluation

This section presents a performance analysis of spherical subsampling in comparison with other range-independent methods (i.e., uniform and random) as well as with the range dependent approach presented in Section 4.4. The targets of the analysis are the three case study scenes introduced in Section 4.2.

5.1. Spatial Distribution

The spatial distribution ratio $P_{sd}$, as computed from (9)-(12), is represented in Fig. 6 for the three scenes with different grid resolutions and two sampling ratios. For all grid sizes, the number of occupied cells without subsampling ($nnz(G_R)$) produces a sparse binary matrix, as few of these cells concentrate a great number of laser points. This is due to the spatial redundancy of the sensor data.

With the same subsampling rate, performance of all methods converges for extreme grid sizes. In general, the range-dependent method shows a good performance, which is only surpassed by the spherical method in the indoor environment for medium and small cell sizes. Independently of the subsampling rate and the grid resolution, the spherical method provides the best spatial distribution index among the range-independent methods. This means that measurement direction equalization avoids loss of relevant scene information by eliminating scan directions that are likely to produce redundant values.

5.2. Matchability After Transformation

Some examples of the application of single spatial transformations (13) to the complete scan $R$ of the flat roof in Fig. 3(b) are shown in Fig. 7. As discussed in Section 3.2, a number of points from the source scan cannot be projected onto the transformed range image. Regions that were not visible from the original point of view appear as white patches, which are clearly visible for all rotations and for the backwards $Y$ translation. White shadows due to a change of perspective are clearly seen for the three translations. These shadows do not appear for the rotations, as the sensor standpoint does not change. In the $Y$ and $Z$ rotations, distinctive curve patterns arise due to the uneven scan density. Note that neither shadows nor resolution artifacts appear for the rotation about $X$, which is the axis of the additional degree of freedom of the 2D scanner.

Figs. 8, 9, and 10 present the value of the $P_{max}$ index for the three scans with $p_t = 0.426$. In each case, the value of this index has been obtained for independent variations of the six possible spatial transformations (in a range of ±8 m for translations and ±40° for rotations). Similar results are obtained with smaller values of $p_t$.

In the case of rotations, matchability depends only on the number of points that remain inside the field of view. Yaw rotation produces almost the same linear reduction of matchable points for all subsampling methods, since this is the additional degree of freedom of the 2D sensor. The spherical method clearly achieves the best matchability for the other two rotations. These results are common for the three scenes.

As for translations, the relative position of the scene objects affects matchability. The abrupt changes about $z_0 = -1$ m correspond to points of view below the ground. Spherical subsampling improves or equals performance for all the transformations except for the translation along the $X$ axis (i.e., sidewise). This is because spherical subsampling eliminates a greater quantity of points from the sides. Nevertheless,
Figure 6: Spatial distribution index of the subsampling methods with respect to grid size for (a) the flat roof, (b) countryside, and (c) building hall.
Figure 7: Examples of projection of the raw flat roof scan with (a) $x_0 = -2$ m, (b) $\gamma = 20^\circ$, (c) $y_0 = -2$ m, (d) $\beta = 20^\circ$, (e) $z_0 = 2$ m, (f) $\alpha = 20^\circ$.

Figure 8: Comparison of subsampling methods with the $P_{mat}$ index for transformations in the flat roof scan.
Figure 9: Comparison of subsampling methods with the $P_{mat}$ index for transformations in the countryside scan.

Figure 10: Comparison of subsampling methods with the $P_{mat}$ index for transformations in the building hall scan.
when using this scan configuration in non-holonomic mobile robots, sidewise translations are smaller than longitudinal translations.

To summarize results, the range-dependent method does not stand out from the range-independent ones, with the exception of the $X$ translation. Uniform and random subsampling achieve similar results in all cases. Finally, the spherical method clearly improves matchability for rotations, as well as for the $Y$ and $Z$ translations.

5.3. Ground Truth Based Performance

The ground truth has been obtained by manually aligning salient features in the case study scenes, starting from the DGPS and IMU estimation. These features can be found naturally in the urban and indoor scenes, but had to be intentionally introduced, as cardboard boxes, in the countryside.

A pair of scans has been obtained from each scene. Thus, two ground truth transformations can be considered for every environment by swapping the roles of the source and destination scans. These resemble forward and backward movement of the sensor in a mobile robot, where the relative $Y$ coordinate ($y_0$) is positive or negative, respectively.

For registration, it is possible to subsample only the source scan or both of them [14]. In the experiments, all the points of the destination scan have been kept. This way, the precision of point matches is improved. Moreover, for the reverse calibration approach, computational cost is not affected.

Tables 1, 2, and 3 show results for the forward and backward transformations in each scene. In these tables, the thresholds have been set to $D = 0.1$ m and $D = 0.05$ m for reverse calibration, and $D = 0.013$ m and $D = 0.0065$ m for closest points. It can be seen that the two indices used for ground truth based performance provide equivalent information, although $P_{rc}$ is always greater than $P_{cp}$ with the selected thresholds. Note that relative performance results are independent of the threshold.

For all cases, the values of $P_{cp}$ and $P_{rc}$ are much higher for the forward movement. In backwards motion, the scanner field of view is opened, so the source scan includes many points from new areas that cannot be matched. All range-independent methods maintain similar rates regardless of the subsampling proportion; however, the performance of the range-dependent approach degenerates for the countryside and the indoor scenes for $p_t = 0.028$. In general, spherical subsampling provides the best rates, and it is unbeaten in forward motion and also in the countryside scene.

5.4. Evaluation of Computational Cost

The computational cost of the data reduction methods implemented in C++ has been measured in an Intel Core 2 Duo processor running at 1833 MHz. Table 4 shows computation times for the obtention of the three range-independent types of binary masks. These only need to be calculated once, which can be done off-line. Building the uniform mask requires less computation time because of its simplicity. The longest times are necessary for the random mask because of random number generation and duplicate avoidance.

Table 5 presents a comparison for online application of a pre-computed range-independent subsampling mask with the average of the range-dependent method for the three environments. This table shows a clear advantage of range-independent subsampling, even if a simple range-dependent approach is considered. In all cases, times are reduced with a smaller subsampling proportion.

6. Conclusions

There are two major problems when spherical scanners are used for real-time scene registration: first, the size of raw scan data imposes a heavy computational cost on point cloud matching; second, certain scene regions are scanned with a higher density, which can distort the registration optimization process. The first limitation can be dealt with by fast range-independent data reduction with random or uniform approaches. However, this solution does not solve the second problem, as data reduction is performed regardless of the scan direction resolutions.

This paper focuses on the analysis of the spherical subsampling method, which was originally proposed to equalize the measure-direction density according to the original sensor resolution [23]. This concept has
Table 1: Ground truth performance indices for different subsampling methods in the flat roof scene.

<table>
<thead>
<tr>
<th>$p_t$</th>
<th>Reduction method</th>
<th>Forward</th>
<th>Backward</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$P_{rc}$ (0.1)</td>
<td>$P_{cp}$ (0.013)</td>
</tr>
<tr>
<td>0.426</td>
<td>spherical</td>
<td>0.719</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.667</td>
<td>0.592</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.666</td>
<td>0.594</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.677</td>
<td>0.556</td>
</tr>
<tr>
<td>0.241</td>
<td>spherical</td>
<td>0.720</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.665</td>
<td>0.595</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.666</td>
<td>0.594</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.650</td>
<td>0.552</td>
</tr>
<tr>
<td>0.108</td>
<td>spherical</td>
<td>0.721</td>
<td>0.631</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.663</td>
<td>0.598</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.666</td>
<td>0.593</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.622</td>
<td>0.521</td>
</tr>
<tr>
<td>0.028</td>
<td>spherical</td>
<td>0.718</td>
<td>0.623</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.665</td>
<td>0.595</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.658</td>
<td>0.581</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.508</td>
<td>0.410</td>
</tr>
</tbody>
</table>

Ground truth: Forward
$x_0 = 0.406 \text{ m}, y_0 = 3.662 \text{ m},$
$z_0 = 0.053 \text{ m}, \alpha = 27.126^\circ,$
$\beta = 5.282^\circ, \gamma = -6.127^\circ$

Ground truth: Backward
$x_0 = -2.017 \text{ m}, y_0 = -3.031 \text{ m},$
$z_0 = -0.566 \text{ m}, \alpha = -27.538^\circ,$
$\beta = -1.879^\circ, \gamma = 7.862^\circ$
Table 2: Ground truth performance indices for different subsampling methods in the countryside scene.

<table>
<thead>
<tr>
<th>$p_t$</th>
<th>Reduction method</th>
<th>$P_{tc}$ (0.1)</th>
<th>$P_{cp}$ (0.013)</th>
<th>$P_{pc}$ (0.05)</th>
<th>$P_{cp}$ (.0065)</th>
<th>$P_{tc}$ (0.1)</th>
<th>$P_{cp}$ (0.013)</th>
<th>$P_{pc}$ (0.05)</th>
<th>$P_{cp}$ (.0065)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.426</td>
<td>spherical</td>
<td>0.573</td>
<td>0.590</td>
<td>0.488</td>
<td>0.483</td>
<td>0.210</td>
<td>0.238</td>
<td>0.176</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.513</td>
<td>0.560</td>
<td>0.426</td>
<td>0.453</td>
<td>0.156</td>
<td>0.180</td>
<td>0.133</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.516</td>
<td>0.562</td>
<td>0.429</td>
<td>0.453</td>
<td>0.156</td>
<td>0.177</td>
<td>0.132</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.492</td>
<td>0.512</td>
<td>0.405</td>
<td>0.393</td>
<td>0.186</td>
<td>0.204</td>
<td>0.151</td>
<td>0.158</td>
</tr>
<tr>
<td>0.241</td>
<td>spherical</td>
<td>0.572</td>
<td>0.587</td>
<td>0.487</td>
<td>0.480</td>
<td>0.208</td>
<td>0.237</td>
<td>0.174</td>
<td>0.195</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.512</td>
<td>0.562</td>
<td>0.429</td>
<td>0.453</td>
<td>0.161</td>
<td>0.174</td>
<td>0.133</td>
<td>0.143</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.508</td>
<td>0.563</td>
<td>0.423</td>
<td>0.454</td>
<td>0.157</td>
<td>0.176</td>
<td>0.130</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.443</td>
<td>0.461</td>
<td>0.369</td>
<td>0.348</td>
<td>0.178</td>
<td>0.203</td>
<td>0.146</td>
<td>0.156</td>
</tr>
<tr>
<td>0.108</td>
<td>spherical</td>
<td>0.572</td>
<td>0.590</td>
<td>0.488</td>
<td>0.480</td>
<td>0.211</td>
<td>0.235</td>
<td>0.178</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.517</td>
<td>0.561</td>
<td>0.422</td>
<td>0.468</td>
<td>0.160</td>
<td>0.174</td>
<td>0.137</td>
<td>0.146</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.511</td>
<td>0.564</td>
<td>0.425</td>
<td>0.456</td>
<td>0.159</td>
<td>0.178</td>
<td>0.133</td>
<td>0.147</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.355</td>
<td>0.382</td>
<td>0.307</td>
<td>0.286</td>
<td>0.181</td>
<td>0.207</td>
<td>0.156</td>
<td>0.159</td>
</tr>
<tr>
<td>0.028</td>
<td>spherical</td>
<td>0.563</td>
<td>0.589</td>
<td>0.483</td>
<td>0.480</td>
<td>0.202</td>
<td>0.222</td>
<td>0.167</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.529</td>
<td>0.562</td>
<td>0.431</td>
<td>0.445</td>
<td>0.152</td>
<td>0.160</td>
<td>0.127</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.521</td>
<td>0.540</td>
<td>0.433</td>
<td>0.446</td>
<td>0.158</td>
<td>0.182</td>
<td>0.136</td>
<td>0.152</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.200</td>
<td>0.160</td>
<td>0.182</td>
<td>0.116</td>
<td>0.149</td>
<td>0.129</td>
<td>0.138</td>
<td>0.092</td>
</tr>
</tbody>
</table>
Table 3: Ground truth performance indices for different subsampling methods in the building hall scene.

$$x_0 = -2.584 \, m, y_0 = 3.135 \, m,$$
$$z_0 = 0.044 \, m, \alpha = 33.333^\circ,$$
$$\beta = -2.714^\circ, \gamma = 5.480^\circ$$

<table>
<thead>
<tr>
<th>$p_t$</th>
<th>Reduction method</th>
<th>$P_{rc}$ (0.1)</th>
<th>$P_{cp}$ (0.013)</th>
<th>$P_{rc}$ (0.05)</th>
<th>$P_{cp}$ (0.0065)</th>
<th>$P_{rc}$ (0.1)</th>
<th>$P_{cp}$ (0.013)</th>
<th>$P_{rc}$ (0.05)</th>
<th>$P_{cp}$ (0.0065)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.426</td>
<td>spherical</td>
<td>0.860</td>
<td>0.704</td>
<td>0.844</td>
<td>0.583</td>
<td>0.353</td>
<td>0.322</td>
<td>0.348</td>
<td>0.277</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.766</td>
<td>0.631</td>
<td>0.754</td>
<td>0.526</td>
<td>0.361</td>
<td>0.330</td>
<td>0.347</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.764</td>
<td>0.623</td>
<td>0.751</td>
<td>0.526</td>
<td>0.361</td>
<td>0.332</td>
<td>0.349</td>
<td>0.295</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.767</td>
<td>0.571</td>
<td>0.752</td>
<td>0.463</td>
<td>0.422</td>
<td>0.363</td>
<td>0.407</td>
<td>0.329</td>
</tr>
<tr>
<td>0.241</td>
<td>spherical</td>
<td>0.861</td>
<td>0.706</td>
<td>0.846</td>
<td>0.584</td>
<td>0.351</td>
<td>0.321</td>
<td>0.346</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.764</td>
<td>0.634</td>
<td>0.755</td>
<td>0.529</td>
<td>0.363</td>
<td>0.326</td>
<td>0.354</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.766</td>
<td>0.627</td>
<td>0.752</td>
<td>0.520</td>
<td>0.361</td>
<td>0.328</td>
<td>0.348</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.750</td>
<td>0.531</td>
<td>0.725</td>
<td>0.431</td>
<td>0.419</td>
<td>0.350</td>
<td>0.417</td>
<td>0.305</td>
</tr>
<tr>
<td>0.108</td>
<td>spherical</td>
<td>0.859</td>
<td>0.703</td>
<td>0.843</td>
<td>0.580</td>
<td>0.349</td>
<td>0.318</td>
<td>0.343</td>
<td>0.270</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.766</td>
<td>0.628</td>
<td>0.753</td>
<td>0.528</td>
<td>0.359</td>
<td>0.326</td>
<td>0.352</td>
<td>0.290</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.767</td>
<td>0.628</td>
<td>0.753</td>
<td>0.524</td>
<td>0.364</td>
<td>0.332</td>
<td>0.353</td>
<td>0.294</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.710</td>
<td>0.482</td>
<td>0.683</td>
<td>0.381</td>
<td>0.457</td>
<td>0.377</td>
<td>0.443</td>
<td>0.314</td>
</tr>
<tr>
<td>0.028</td>
<td>spherical</td>
<td>0.853</td>
<td>0.701</td>
<td>0.840</td>
<td>0.579</td>
<td>0.347</td>
<td>0.326</td>
<td>0.341</td>
<td>0.271</td>
</tr>
<tr>
<td></td>
<td>random</td>
<td>0.772</td>
<td>0.632</td>
<td>0.770</td>
<td>0.521</td>
<td>0.356</td>
<td>0.333</td>
<td>0.347</td>
<td>0.301</td>
</tr>
<tr>
<td></td>
<td>uniform</td>
<td>0.773</td>
<td>0.638</td>
<td>0.763</td>
<td>0.528</td>
<td>0.366</td>
<td>0.335</td>
<td>0.352</td>
<td>0.301</td>
</tr>
<tr>
<td></td>
<td>range</td>
<td>0.460</td>
<td>0.331</td>
<td>0.447</td>
<td>0.250</td>
<td>0.361</td>
<td>0.291</td>
<td>0.353</td>
<td>0.213</td>
</tr>
</tbody>
</table>

Table 4: Generation times for the range-independent masks.

<table>
<thead>
<tr>
<th>Subsampling</th>
<th>$p_t = 0.426$</th>
<th>$p_t = 0.241$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spherical</td>
<td>3.08 ms</td>
<td>1.13 ms</td>
</tr>
<tr>
<td>Random</td>
<td>8.43 ms</td>
<td>1.50 ms</td>
</tr>
<tr>
<td>Uniform</td>
<td>0.75 ms</td>
<td>0.21 ms</td>
</tr>
</tbody>
</table>

Table 5: Subsampling times for the range-dependent and independent methods.

<table>
<thead>
<tr>
<th>Subsampling</th>
<th>$p_t = 0.426$</th>
<th>$p_t = 0.241$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range-independent</td>
<td>0.76 ms</td>
<td>0.54 ms</td>
</tr>
<tr>
<td>Range-dependent</td>
<td>14.69 ms</td>
<td>11.56 ms</td>
</tr>
</tbody>
</table>
been extended to allow for further data reduction. The main advantage of spherical subsampling is that it considers the scanning procedure for scene perception, so it is less likely that relevant scene information is lost. This approach has the same low computational cost as other range-independent methods, as uniform and random subsampling, because it can be applied as a pre-computed binary mask.

Furthermore, the paper introduces three types of performance criteria to evaluate subsampling methods intrinsically, independently of a given registration technique. Spatial distribution and matchability after transformation are indices to measure the conservation of scene information that can be computed from a single scan, so they do not rely on the ground truth. Moreover, if the ground truth is known, matching performance can be evaluated by the number of point matches for a couple of scans with both the closest points and the reverse calibration indices.

An experimental analysis of fast subsampling methods has been performed with a customized 3D scanning system and three representative scenes. Spherical subsampling has been compared with the other range-independent methods and with a simple range-dependent data reduction approach. Analysis of experimental data shows that spherical subsampling generally achieves a better performance than uniform and random methods. This results show that spherical subsampling is more efficient to eliminate redundant information. Moreover, this approach exhibits comparable results with a more computationally expensive range-dependent subsampling method.

Future work includes the use of spherical subsampling as a preprocessing step in combination with specific range-dependent techniques.

References


