



DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

3D MAPPING FROM SPARSE LIDAR DATA

3D MAPOVÁNÍ S VYUŽITÍM ŘÍDKÝCH DAT SENZORU LIDAR

PH.D. THESIS
DISERTAČNÍ PRÁCE

AUTHOR
AUTOR PRÁCE

ING. MARTIN VELAS

SUPERVISOR
VEDOUCÍ PRÁCE

PROF. ING. ADAM HEROUT, PH.D.

BRNO 2019

ABSTRACT

This work presents a human-carried mapping backpack based on a pair of Velodyne LiDAR scanners. The system is a universal solution for both large scale outdoor and smaller indoor environments. It benefits from a combination of two LiDAR scanners, which makes the odometry estimation more precise. The scanners are mounted under different angles, thus a larger space around the backpack is scanned. By fusion with GNSS/INS sub-system, the mapping of featureless environments and the georeferencing of resulting point cloud is possible. By deploying SoA methods for registration and the loop closure optimization, it provides sufficient precision for many applications in BIM (Building Information Modeling), inventory check, construction planning, etc. In our indoor experiments, we evaluated our proposed backpack against ZEB-1 solution, using FARO terrestrial scanner as the reference, yielding similar results in terms of precision, while our system provides higher data density, laser intensity readings, and scalability for large environments.

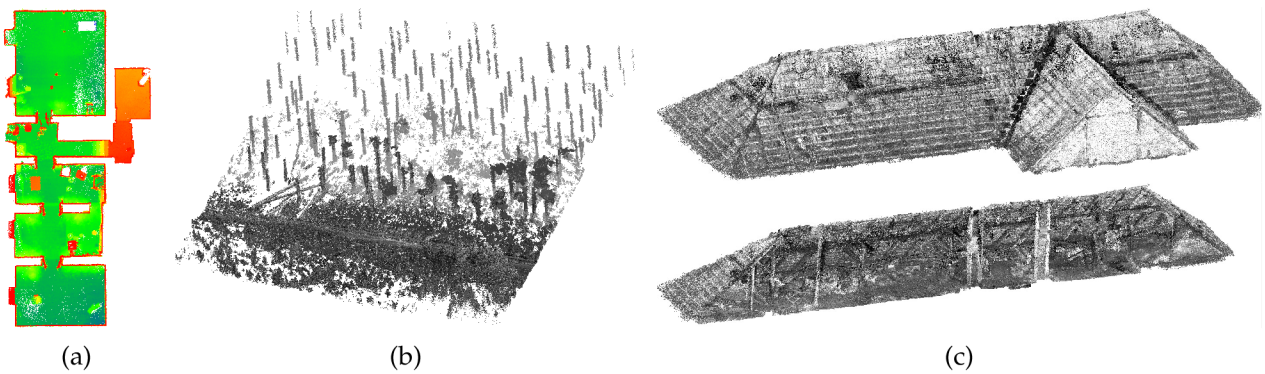


Figure 0.1: The motivation and the results of our work. The reconstruction of indoor environments (a) is beneficial for inspection, inventory checking and automatic floor plans generation. 3D maps of forest environments (b) is useful for quick and precise estimation of the biomass (timber) amount. The other example of 3D LiDAR mapping deployment is preserving cultural heritages or providing models of historical building, e.g., the roof in (c).

INTRODUCTION

In recent years, the LiDAR (Light Detection And Ranging) technology has become very popular in the field of geodesy and related fields, where the availability of 3D models of outdoor or indoor environments can be beneficial: e.g., forestry, architecture, preserving cultural heritage, construction monitoring, etc. The examples of reconstructions from similar practical applications can be found in Fig. 0.1. Using 3D mapping can also be beneficial for time and cost reduction. The same model can be shared among different professionals in different fields of expertise without the need for personal inspection and measuring at a given place individually.

This demand causes a huge interest in developing solutions that would be able to capture the reality and provide reliable 3D reconstructions out of the box. However, there are also other requirements for such a system.

The data acquisition process has to be quick and the planning of fieldwork should be minimized. This requirement discriminates solutions based on static terrestrial lasers (e.g., Leica and Riegl of FARO companies), requiring detailed planning of the data acquisition and manual system set up on a tripod within multiple convenient viewpoints across the scene.

The solution has to be mobile and easy to handle. This naturally leads to the preference of human carried (backpack or handheld) solutions instead of terrestrial or vehicle based solutions, such as NavVis ¹, which, for example, does not support traversing tilted surfaces such as ramps.

However, the necessity for reliability in terms of resulting model precision is in contradiction with these two requirements. Stationary terrestrial LiDAR solutions require time demanding scanning process while providing a great accuracy (in order of millimeters) because of fewer degrees of freedom. Although, for many applications listed above, there is no need for such precision, our goal is the difference between the reality and the resulting 3D model below 5 cm. This value was requested by the experts in the field of geodesy with whom we consulted.

In the practical applications, completeness of the final map should also be guaranteed because it might be difficult to repeat the scanning. The operator has to be aware of the fact that all necessary data of the whole environment were acquired. We fulfilled this requirement by providing a live preview of the collected data.

The resulting model has to be dense enough, so that all important objects such as furniture and other inventory can be recognized and distinguished. This is the typical issue of existing solutions such as ZEB-1, where no LiDAR intensity readings are available. Therefore, our solution relies on Velodyne LiDARs, which provide a huge amount of data and the resulting models are dense (see examples in Fig. 1.1). It also provides the laser intensity readings, which do not depend on the lighting conditions, contrary to camera-aided solutions. Moreover, we propose laser intensity normalization, which increases the recognizability of the objects since the laser inten-

¹ <https://www.navvis.com>

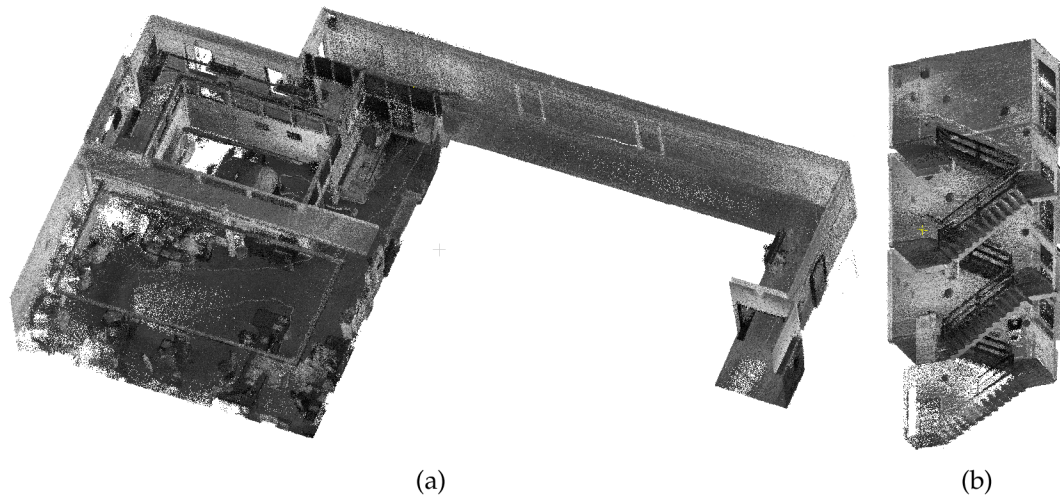


Figure 1.1: The example of resulting models of indoor mapping. The office environment (a) and the staircase (b) were captured by a human carrying our 4RECON backpack. The data acquisition process took 3 and 2 min, respectively.

sity readings cannot be considered as the “color” of the object as it depends on the range of measurement, the angle of incidence, and the emitted energy.

Some of the existing solutions are not comfortable enough to use. According to practical experience of the operators, handheld solutions such as ZEB are physically difficult to operate for a longer period of time since the mapping head weighs approximately 0.4–1 kg, and it has to be carried or swept by hand.

The final requirement is an affordable price. We use Velodyne VLP-16 scanners, which are relatively cheap in comparison to the other LiDAR solutions, and a universal IMU (Inertial Measurement Unit) solution, which can be upgraded by a dual antenna and therefore reused in the outdoor environment where GNSS (Global Navigation Satellite System) is available.

The contributions of this work can be summarized as the proposal of a LiDAR mapping solution with the following characteristics:

- It is capable of both small indoor and large open outdoor environments mapping, georeferencing and sufficient precision in the order of centimeters. These abilities are evaluated using multiple datasets.
- It benefits from a synchronized and calibrated dual LiDAR scanner, which significantly increases field of view. Both scanners are used for both odometry estimation and 3D model reconstruction, which enables scanning of small environments, narrow corridors, staircases, etc.
- It provides the ability to recognize objects in the map due to sufficient point density and our novel intensity normalization for the measurements from an arbitrary range.

We also performed a precise evaluation and comparison of our previously proposed point cloud registration method CLS (Collar Line Segments) with state-of-the-art approach LOAM (LiDAR Odometry and Mapping), which has not yet been published. Moreover, we upgraded our CLS method with automatic overlap estimation for better registration flexibility.

This section consists of two main parts: First, the hardware design concepts are introduced. Then, the software solutions dealing with calibration, precise odometry estimation, alignment and intensity normalization are presented.

The design of the solution follows the requirements elaborated in Sec. 1. They have been carefully formulated and discussed with experts in the field of geodesy and geospatial data processing. Besides the essential goal of reliable 3D reconstruction performed automatically, which is demonstrated in the following section, the proposed solution does the following:

- It fulfils the requirements for precision of the model up to 5 cm. Thanks to the robust loop closure, ambiguities (e.g., “double wall” effects) are avoided.
- The system is comfortable to use and it is as mobile as possible. The backpack weighs 9 kg (plus 1.4 kg for the optional dual antenna extension), and it is easy to carry around various environments including stairs, narrow corridors, rugged terrain, etc.
- The pair of synchronized and calibrated Velodyne LiDARS increases the field of view (FOV) and enables mapping of small rooms, narrow corridors, staircases, etc. (see Fig. 2.1) without the need for special guidelines for scanning process.
- The data acquisition process is fast with verification of data completeness. There are no special guidelines for the scanning process (comparing to the requirements of ZEB) and the operator is required only to visit all places to be captured in a normal pace. Moreover, captured data are visualized online at the mobile device (smartphone, tablet) for operator to see whether everything is captured correctly.
- Since we are using long range Velodyne LiDAR (compared to simple 2D rangefinders such as Hokuyko or Sick) and optional GNSS support, we provide a universal economically convenient solution for both indoor and outdoor use. For such scenarios, where GNSS is available, final reconstruction is georeferenced—the 3D position in the global geographical frame is assigned to every 3D point in the model.
- The final 3D model is dense and colored by the laser intensity, which is further normalized. This helps distinguishing important objects, inventory, larger texts, signs, and some surface texture properties.

2.1 HARDWARE DESCRIPTION

The core of our backpack, in Fig. 2.2, is the pair of Velodyne LiDAR¹ scanners VLP-16 (Pucks). Each of them contains 16 laser transmitter–receiver pairs, which are reflected into the environment by a rotating mirror with 10 Hz frequency. This frequency can be decreased or increased up to 20 Hz. However, frequency higher than 10 Hz causes se-

¹ <https://velodynelidar.com/>

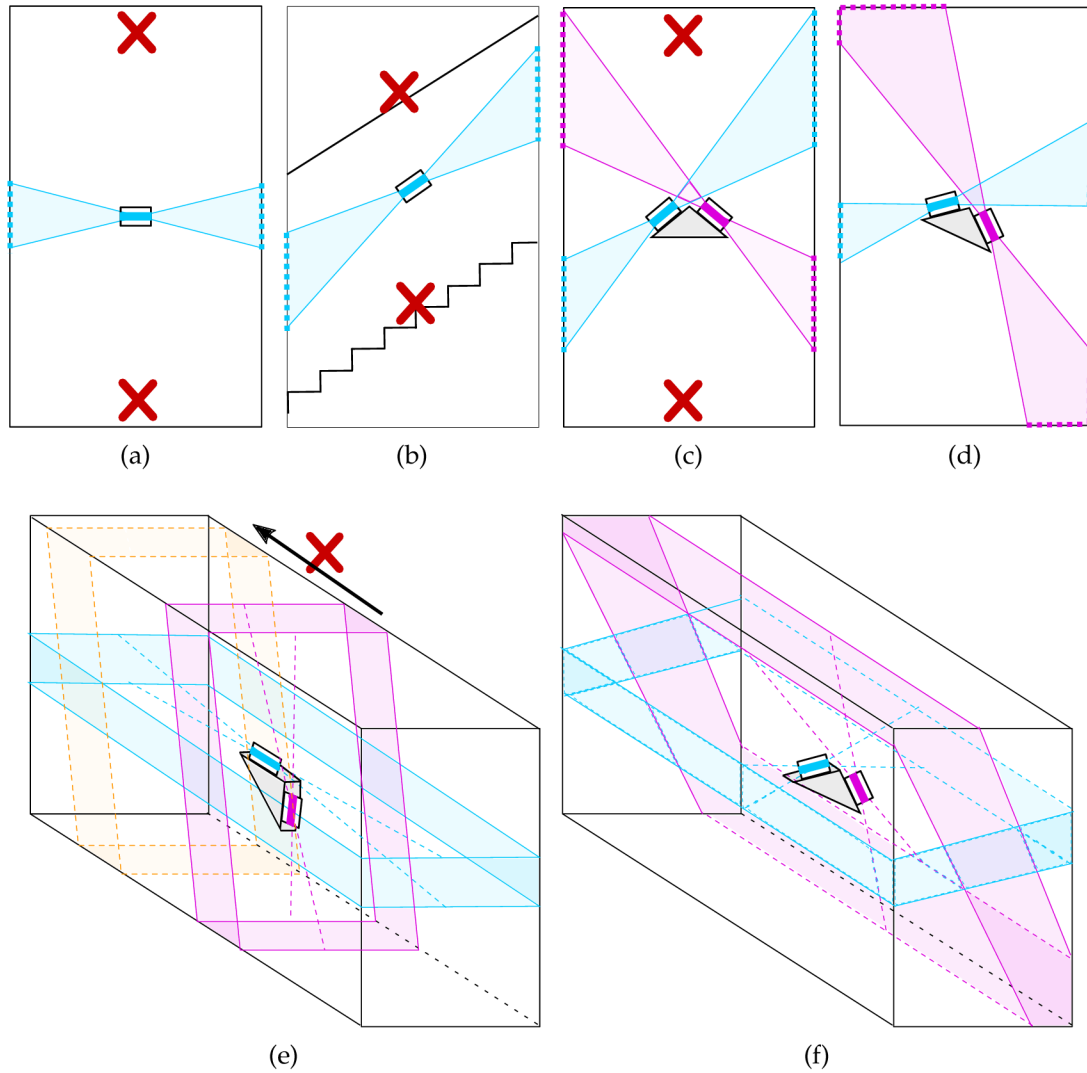


Figure 2.1: Various configurations of LiDAR scanners in worst case scenarios we have encountered in the experiments: narrow corridor (a),(c) and staircase (b). The field of view (30° for Velodyne Puck) is displayed in color. When only single LiDAR (a) was used, the scans did not contain 3D information of the floor or the ceiling (red cross). The situation was not improved when the scanner is tilted because of failing in, e.g., staircases (b). When we added a second LiDAR, tiled asymmetrical configuration (d) provides better top-bottom and left-right observation than the symmetrical one (c). Moreover, when the LiDARs are aligned in direction of movement (e), there is no overlap between current (violet) and future (yellow) frame, leading to lower accuracy. In the solution (f), the LiDARs are aligned perpendicularly to the walking direction solving all mentioned issues..

rious undesirable vibration of the sensor, which makes precise odometry estimation impossible. The rotation gives the sensor 360° horizontal FOV with 0.2° horizontal resolution. Vertically, the laser beams are evenly distributed with 2° resolution covering 30° vertical FOV. Each of the scanners weighs 830 g and is considered to be a hybrid solid state LiDAR, since there are no outer moving parts. This type of scanner is able to reach 100 m range with precision around 2 cm. As mentioned above, Velodyne scanners provide also values of intensity readings, which corresponds to the surface reflectivity.

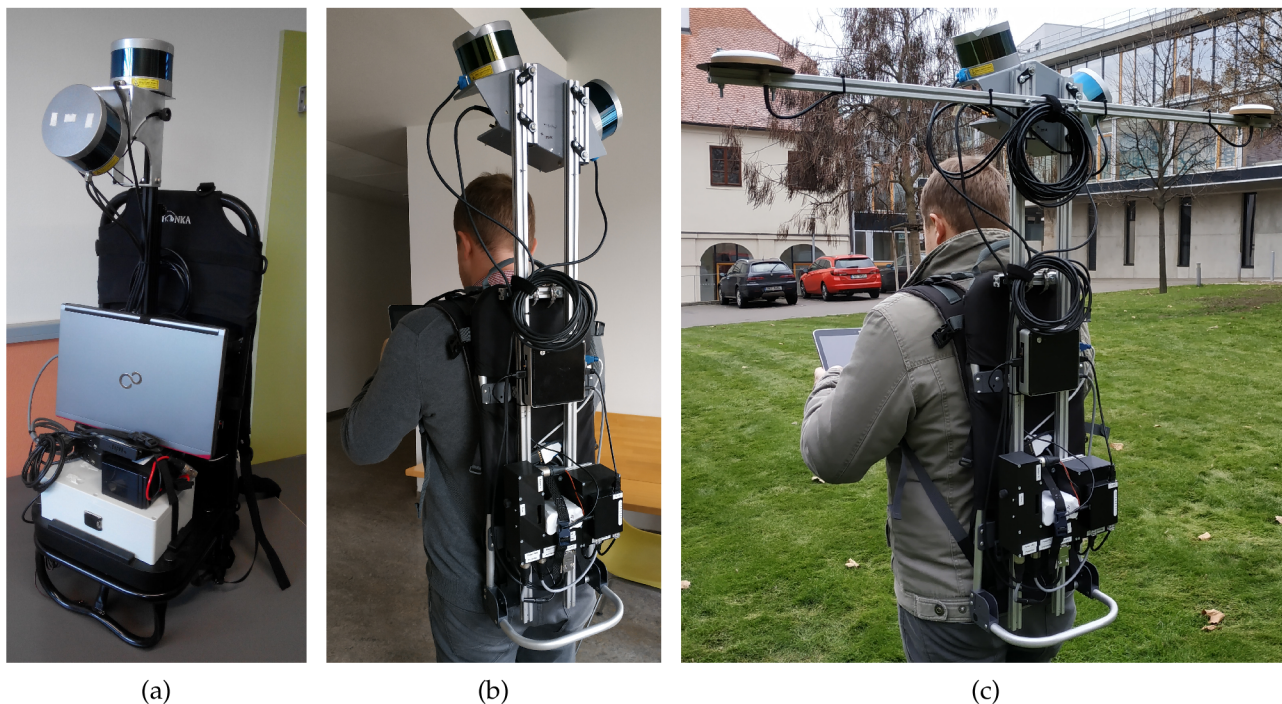


Figure 2.2: The initial (a) and improved (b),(c) prototype of our backpack mapping solution for both indoor (b) and outdoor (c) use. The removable dual GNSS antenna provides precise heading information, aiding for outdoor odometry estimation and also georeferencing of the resulting 3D point cloud model. It should be noted that the position of LiDAR scanners is different in the initial and the later solution. This is elaborated on in the next section.

As the aiding sensor, the GNSS/INS (Inertial Navigation System) Advanced Navigation SpatialDual² is deployed. It integrates multiple sensors such as accelerometers, gyroscopes, magnetometer, pressure sensor, and most importantly—the dual-antenna GNSS subsystem providing reliable heading information. With RTK (Real Time Kinematics) or PPK (Post-Processed Kinematics) corrections, the system should provide 8 mm horizontal and 15 mm vertical positional accuracy, and 0.03° and 0.06° orientation precision in terms of roll/pitch and heading angle, respectively. Precise heading information is provided by a dual antenna solution and therefore it is only available outdoors. This limitation also holds for positional data. For indoor scenarios, only roll and pitch angles are reliable and they are relevant for horizontal alignment. The unit weighs 285 g and besides the 6 DoF (six Degrees of Freedom including 3D position and rotation) pose estimation it also provides 1PPS (Pulse Per Second) and NMEA messages for precise synchronization of both Velodyne LiDAR scanners. The details regarding wiring the components can be found in Fig. 2.3.

The rest of the hardware is responsible for controlling the data acquisition and storing the data (Intel NUC Mini PC), and powering all the components with small Li-Ion battery with capacity 10,400 mAh lasting approximately 2 h.

² <https://www.advancednavigation.com/product/spatial-dual>

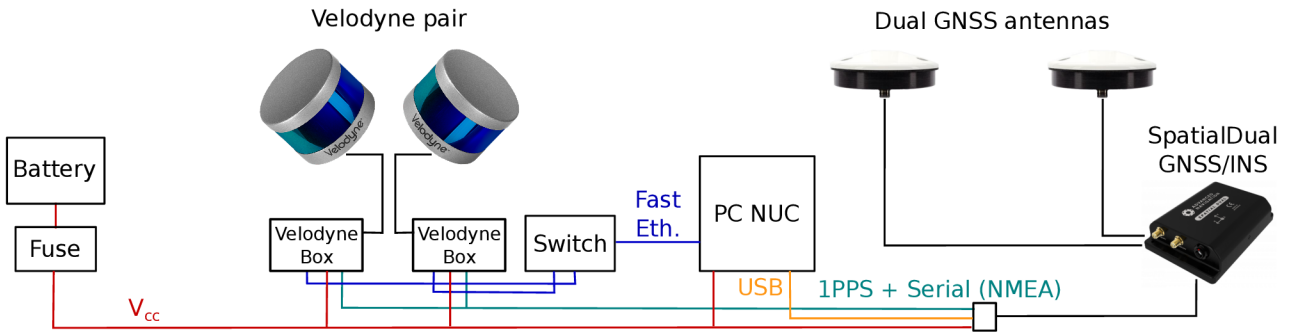


Figure 2.3: Components of the system and the connections. Each Velodyne scanner is connected via a custom wiring “box” requiring power supply (red wires), 1PPS and NMEA synchronization (green) and Fast Ethernet (blue) connection with computer (PC NUC in our case).

2.2 DUAL LIDAR SYSTEM

During the experiments, we discovered that the limited (30°) horizontal field of view is not an issue for large open spaces. However, when the space is getting smaller and the environment shrinks (e.g., corridors narrower than 2 m), such a field of view causes serious problems, leading to poor accuracy or even total failures of the SLAM system. The worst cases and our solutions are displayed in Fig. 2.1. We experimentally discovered that we need at least two synchronized Velodyne Puck scanners to provide a robust solution that covers both the floor/ceiling and the walls, even in small or narrow rooms.

To achieve good accuracy and to cover the environment, the scanners are mounted perpendicular to the direction of the operator movement—one in horizontal and second in vertical orientation, as displayed in Figures 2.2b,c and 2.1f. All other configurations (e.g., Configuration e.) in our initial prototype in Fig. 2.2a were not able to capture both horizontal and vertical properties of the environment, or did not provide a large coverage necessary for precise pose estimation.

2.3 CALIBRATION OF THE SENSORS

To leverage the full potential of using two Velodyne LiDARs, these scanners have to be properly synchronized and calibrated. As mentioned above, the sensors are synchronized via NMEA messages (GPS communication protocol) and 1PPS (Pulse Per Second) signal provided by SpatialDual inertial navigation system. Sufficient intrinsic calibration parameters of LiDAR scanners themselves (corrections) are provided by Velodyne company and processed by the driver (in ROS Velodyne package).

Therefore, the task to solve is the estimation of extrinsic calibration parameters in terms of relative 6DoF pose estimation for both laser scanners C_{V1} , C_{V2} and INS sensor C_I in Fig. 2.4. First, the transformation between the scanners is computed. To do so, two 3D maps of a large indoor space (a large lecture hall in our case) were built by the scanners separately using our previously published method [16]. These two 3D maps are ICP aligned. The resulting 3D geometrical transformation represents mutual position of the sensors $C_{V1}^{-1} * C_{V2}$ and also the alignment of laser data they provide

as presented in Fig. 2.5. Since we are interested only in relative transformations between the sensors, the origin can be arbitrarily defined, e.g., as the position of the first Velodyne and $\mathbf{C}_{V1} = \mathbf{I}$. A single frame point cloud consists of multiple (two in our case) synchronized LiDAR frames and therefore it will be denoted as the *multiframe*.

To be able to use data provided by the INS system, an extrinsic calibration \mathbf{C}_I between the laser scanners and the INS sensor needs to be estimated. All sensors are fixed on the custom made aluminum mount and therefore the translation parameters can be found in the blueprints of the mount or can be measured with millimeter precision. However, mutual rotation has to be estimated more precisely, because just a fraction of degree misalignment would cause serious errors for long range laser measurements.

We found that the rotation parameters as the transformation between the floor normal vector \vec{n}_i in the point cloud data and the gravity vector \vec{g}_i provided by the INS sensor, since these vectors should be aligned. Points of the floor are selected manually and the normal of the best fitting plane is computed. This can be performed in arbitrary software for visualization and processing of the point clouds—CloudCompare³ in our case. We performed multiple measurements for different inclines of the backpack in the indoor corridor with a perfectly straight floor. The final rotation \mathbf{R}_{C_I} between the Velodynes and INS sensor was estimated by SVD (Singular value Decomposition) [11] (Equation (2.2)) of covariance matrix \mathbf{A} of these 3D vector pairs (Equation (2.1)) (floor normal and the gravity). Multiplication with matrix \mathbf{E} (Equation (2.5)) solves the ambiguity between right/left hand rotation—we always compute right-hand representation. Equations (2.1)–(2.5) are based on the work [11].

$$\mathbf{A} = \sum_i \vec{n}_i^T \cdot \vec{g}_i \quad (2.1)$$

$$\mathbf{U}\mathbf{\Sigma}\mathbf{V}^* = \mathbf{A} \quad (2.2)$$

$$e = \begin{cases} 1, & \text{if } |\mathbf{V}\mathbf{U}^T| \geq 0 \\ -1, & \text{otherwise} \end{cases} \quad (2.3)$$

$$\mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e \end{bmatrix} \quad (2.4)$$

$$\mathbf{R}_{C_I} = \mathbf{V}\mathbf{E}\mathbf{U} \quad (2.5)$$

2.4 POINT CLOUD REGISTRATION

The core element of the software part is the alignment of the point cloud data into a 3D map of the environment. There are multiple state-of-the-art approaches for point cloud registration and odometry estimation, including our previously published approach Collar Line Segments [16]. We compared our approach with LOAM [17] algorithm, using the implementation available. The results of this experiment are pre-

³ <https://www.danielgm.net/cc/>

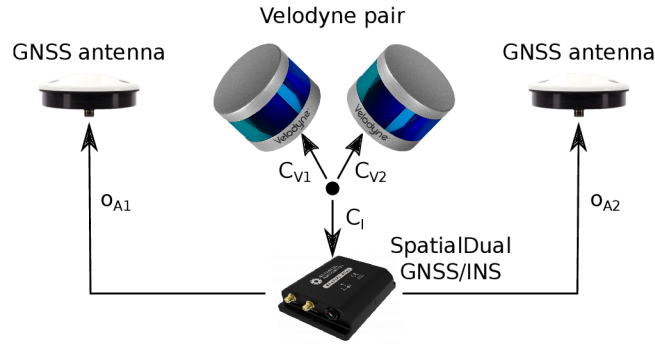


Figure 2.4: Extrinsic calibration required in our system. The mutual positions between the Velodyne scanners and the GNSS/INS unit are computed. The offsets \vec{o}_{A1} , \vec{o}_{A2} of the antennas are tape measured.

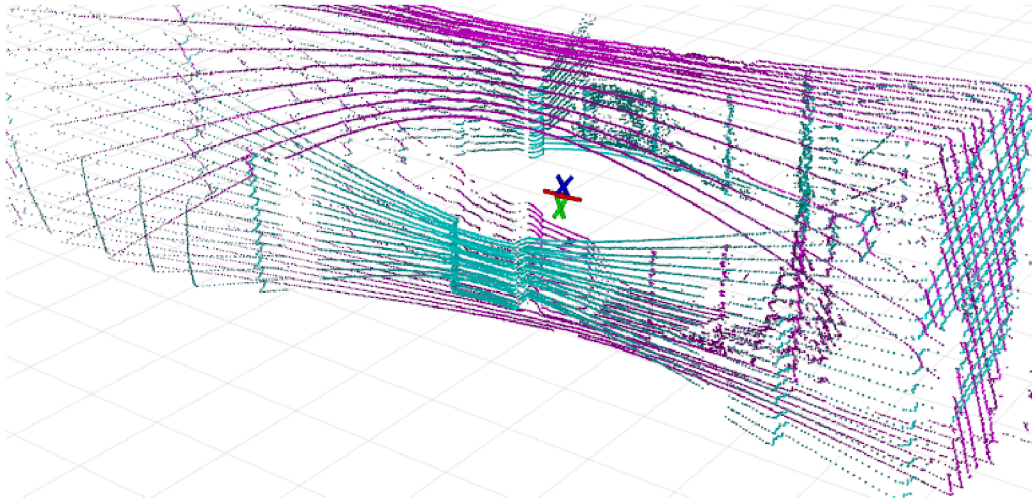


Figure 2.5: Two Velodyne LiDAR frames aligned into the single *multiframe*. This data association requires time synchronization and precise extrinsic calibration of laser scanners.

sented in Table 3.1, which shows the superior accuracy of our method, thus CLS was a natural choice for our mapping backpack solution.

The basic idea of the CLS method is to overcome the data sparsity of 3D LiDAR scanner (e.g., Velodyne) by sampling the data by line segments. The points captured by individual laser beams form so called “ring” structures displayed in Fig. 2.6a. There is a large empty space between these rings and while moving, same places of the scene are not repeatedly scanned, valid matches are missing and the closest point approaches (e.g., ICP) are not applicable. By using CLS, the space between the rings is also covered and correct matching of structures in the LiDAR frames is enabled.

The environment in the field of view is represented by the set of CLS line segments. They are randomly generated between the neighboring ring points within the azimuthal bin as described in Fig. 2.6a. Since we are using two LiDAR scanners, collar line segments are generated for the scans of each sensor individually. Using the transformation established by extrinsic calibration described in Sec. 2.3, line segments are transformed and joined into the single set for each multiframe.

After the sampling is done, matching of the closest line segments is performed. The line segments are extended into the infinite lines, and the closest points between matching lines are used for direct estimation of translation. SVD [11] is used again for

estimation of rotation parameters in the same manner, as described in Sec. 2.3. This description is only a brief introduction to the CLS method and more information can be found in our previous publication [16].

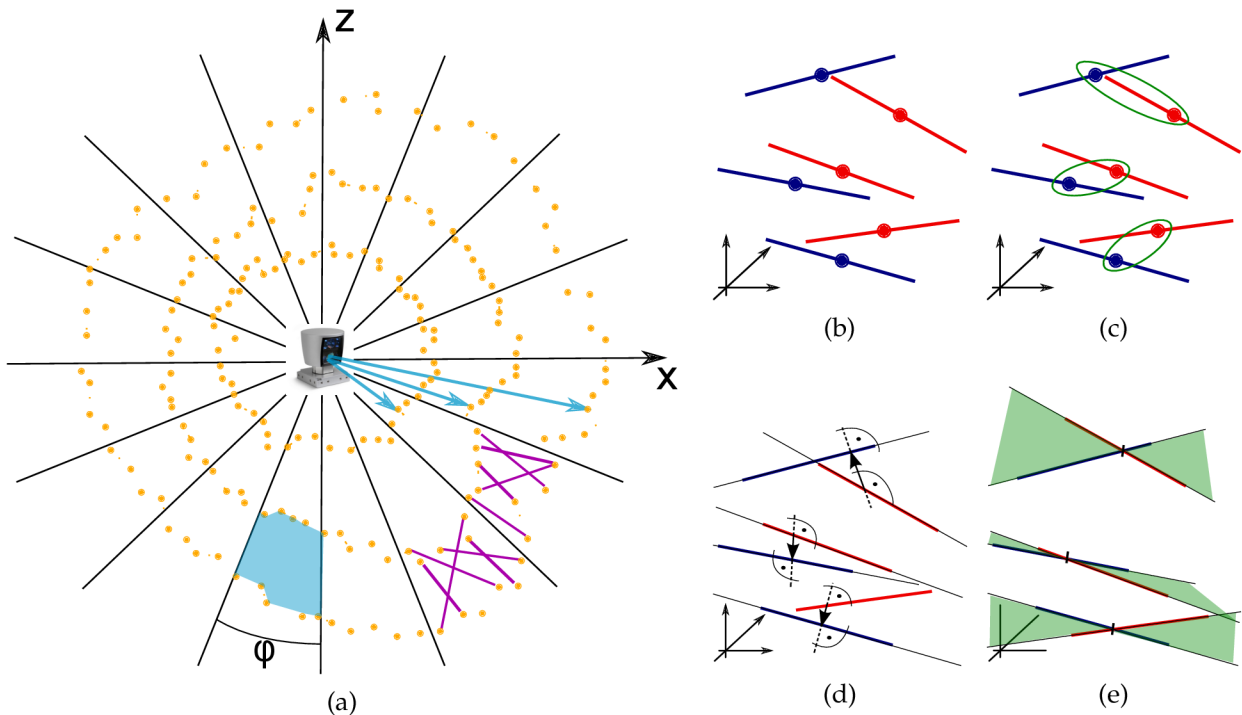


Figure 2.6: The sampling of Velodyne point cloud by the Collar Line Segments (CLS) (a). The segments (purple) are randomly generated within the polar bin (blue polygon) of azimuthal resolution ϕ . The registration process (b–e) transforms the line segments of the target point cloud (red lines) to fit the lines of the source cloud (blue). First, the lines are matched by Euclidean distance of midpoints (c); then, the segments are extended into infinite lines and the vectors between closest points are found (d); and, finally, they are used to estimate the transformation that fits the matching lines into common planes (green in (e)).

2.5 OVERLAP ESTIMATION

This work provides a novel solution for automatic estimation of the core parameter of the CLS approach. Before the transformation is estimated, invalid matches must be discarded. In our previous work, this was done by a simple distance thresholding, or by keeping a certain portion of matches (e.g., 50%). However, using a constant threshold or portion value is not flexible enough. It can cause significant registration misalignments, when invalid matches are used, or insufficient convergence when the valid matches are ignored.

Assuming that an initial coarse alignment is known, we are able to estimate *the overlap* between these frames and use this value as the portion of matches to keep (e.g., for 30% overlap, 30% of best matches are kept). This solution adapts to the specific situation of each pair of LiDAR frames to be registered and leads to a significantly better precision.

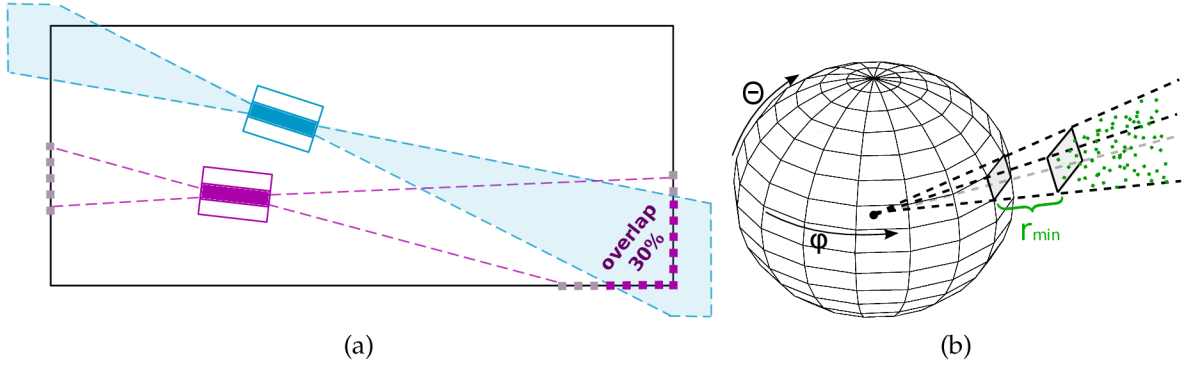


Figure 2.7: The overlap (a) between the source (blue) and the target (purple) LiDAR frame. In this case, approximately 30% of source points are within the view volume of target frame. The view volume can be effectively represented by *spherical z-buffer* (b) where range information (minimum in this case) or the information regarding empty space within the spherical grid is stored.

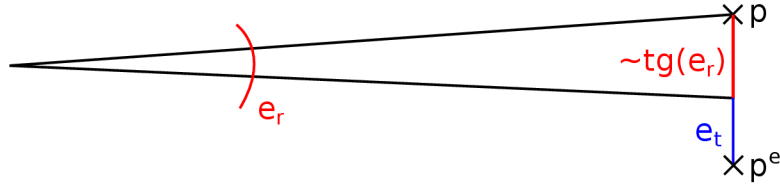


Figure 2.8: The error of measurement (Euclidean distance between points p and p^e) can be split into rotation e^r and translation e^t part. The impact of rotation error $2 \cdot \text{tg}(e_r/2)$ can be simplified to $\text{tg}(e_r)$ due to near linear properties of tangent function for small angles.

The overlap value (Fig. 2.7a) is effectively estimated by *spherical z-buffer* structure [15] in Fig. 2.7b. First, the target cloud is transformed into the source cloud coordinate frame and the $[x, y, z]$ coordinates of all the points are transformed to spherical coordinates ϕ, θ, r (polar angle, elevation angle, and range). Each spherical bin of the z-buffer is assigned with minimal range value from the source point cloud. The minimal value is chosen since unwanted reflections sometimes cause invalid long range measurements and therefore there is the best chance that the minimum range measurement is valid. Then, all the points of target point cloud (also transformed to spherical coordinates) with range below the value in z-buffer (including certain tolerance) are considered to be overlapping points and the ratio to all the points is considered to be the *overlap* value. More formally, if the point p with range p_r within the spherical bin i fulfills the requirement

$$p_r < r_{\min}^i \cdot t_r + t_a, \quad (2.6)$$

it is considered to be a part of the overlap. Value r_{\min}^i denotes the minimal range value stored within the spherical bin. Absolute t_a and relative t_r tolerance values represent the acceptable translation and rotation error. Especially the error of rotation causes larger displacements for larger ranges. Equation (2.6) follows our *error model*, where the error e is the distance between precise point coordinates p (which are unknown) and known erroneous coordinates p^e which can be approximately estimated as:

$$e = |p - p^e| = p_r^e \cdot \text{tg}(e_r) + e_t, \quad (2.7)$$

where e_r represents rotation, e_t is the translation error, and p_r^e is the range of the erroneous point (see also Fig. 2.8). In our experiments, we used the tolerance values $t_r = 0.1$ and $t_a = 0.3$ for the overlap estimation. This allows rotation error e_r approximately 5° and translation error 30 cm for the initial coarse transformation between the scans.

2.6 ROLLING SHUTTER CORRECTIONS

As mentioned in the description of Velodyne sensor, spinning frequency is approximately 10 Hz which leads to 100 ms duration of a single LiDAR scan acquisition. This is a relatively long time when significant movement is assumed. Large translation in the case of fast vehicles or possible fast rotations in case of human carrier can cause distortions in LiDAR frame displayed in Fig. 2.9. We denote this effect as *rolling shutter* because it resembles rolling shutter distortion of optical sensors.

This means that the LiDAR data cannot only be rigidly transformed, but a continuous transformation needs to be applied or at least approximated. The single Velodyne Puck frame consists of approximately 75 packets, each carrying a slice of the frame. Slices are evenly distributed in both time and space. Thus, for each i th frame, we compute the relative transformation $\mathbf{T}_{i \rightarrow j}$ that occurred during the acquisition of the current frame using the global position \mathbf{P}_i of the current frame and the pose \mathbf{P}_{i+1} of the next one as:

$$\mathbf{T}_{i \rightarrow j} = \mathbf{P}_i^{-1} \cdot \mathbf{P}_{i+1}. \quad (2.8)$$

The correction for each slice is estimated by interpolation of this transformation. The translation parts are interpolated linearly and, for the rotations, Spherical Linear Interpolation (SLERP) [13] over quaternion representation is used. For the first slice, zero transformation is estimated and the last one is transformed by $\mathbf{T}_{i \rightarrow j}$.

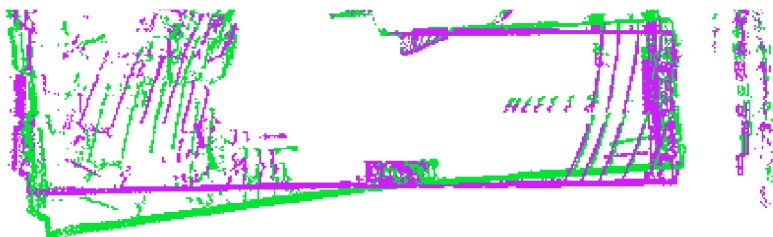


Figure 2.9: Example of a LiDAR frame distorted by the rolling shutter effect when the operator with mapping backpack was turning around (green) and the corrected frame (purple). This is the top view and the distortion is mostly visible on the “bent” green wall at the bottom of this picture.

2.7 POSE GRAPH CONSTRUCTION AND OPTIMIZATION

The proposed CLS method for point cloud alignment can only provide consecutive frame-to-frame registration. However, since each registration is burdened by a small error, after some time, the accumulated error (*drift*) is no longer acceptable. To reduce this drift and also to close loops of revisited places, we propose an iterative process of *progressive pose graph* construction and optimization. The key idea of this algorithm

is progressive refinement of odometry estimation from local precision within small time window to global precision across the whole model. This iterative method is described in Fig. 2.10 and more formally in Algorithm 1.

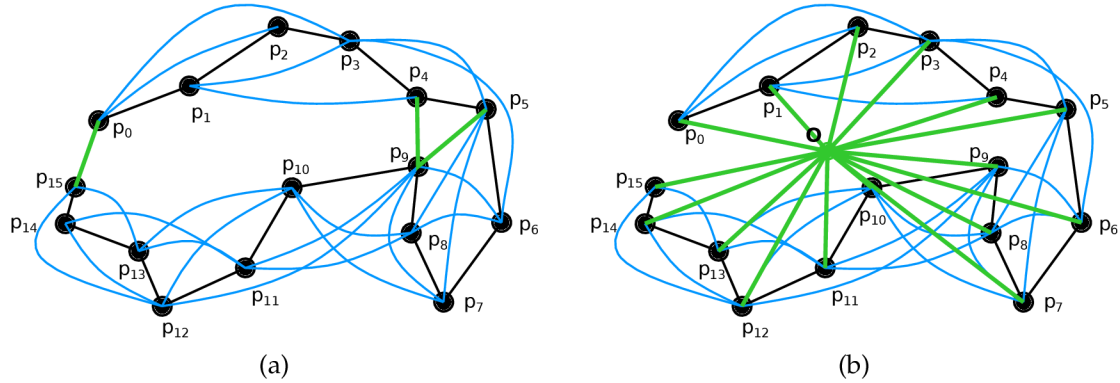


Figure 2.10: Pose graph as the output of point cloud registration and the input of SLAM optimization. The goal is to estimate 6DoF poses $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N$ of graph nodes (vertices) p_1, p_2, \dots, p_{15} in the trajectory. The edges represent the transformations between LiDAR frames for given nodes estimated by point cloud registration. Black edges represent transformations between consequent frames, blue edges are for transformations within a certain neighborhood (maximum distance of three frames in this example) and the green edges (in (a)) represent visual loops of revisited places detected by a significant overlap between the given frames. When GNSS subsystem is available (b), additional visual loops are introduced as transformations from the origin \mathbf{O} of some local geodetic (orthogonal NED) coordinate frame.

First, only consecutive frames (within neighborhood of size 1) are registered, and then the neighborhood is gradually enlarged (size d in Algorithm 1, step 1) until it covers all N frames. CLS registration is performed for each pair (i th and j th frame) within the current neighborhood where a significant overlap is found and then efficient pose graph optimization using SLAM++ framework [4] is performed. Modulo operator in Step 3 reflects the fact that we assume a circular trajectory. This assumption of beginning and ending the data acquisition process at the same place is common also for other similar solutions (ZEB-1, ZEB-REVO, etc.) [3]. It helps the system to identify at least one visual loop that guarantees reasonable results from the global SLAM-based optimization.

Before a pair of frames is registered, the presence of overlap larger than t_o is verified (Line 5 in Algorithm 1) in order to preserve the registration stability. We used minimal 0.5 overlap in our experiments. This also plays the role of visual loop detection every time a place is revisited.

Moreover, after the CLS registration is performed, we verify the result of registration (Line 8) using the error model described in Equation (2.7). As the reference range value, we take the median range of the source point cloud. In our experiments, we used tolerance values $t_r = 0.01$ and $t_a = 0.05$ representing tolerance of approximately 0.5° in rotation and 5 cm in positional error.

For outdoor mapping, the absolute position and orientation are provided by the GNSS/INS subsystem with PPK (Post Processed Kinematics) corrections. While the global error of these poses is small, relative frame-to-frame error is much larger when compared to the accuracy of pure SLAM solution. Therefore, we combine our SLAM

Algorithm 2.1: Progressive refinement of 6DoF poses $\{\mathbf{P}_i\}_{i=1}^N$ for sequence of frames $\{\mathbf{f}_i\}_{i=1}^N$ by optimizing pose graph \mathbf{G} .

```

1: for  $d = 2$  to  $\frac{N}{2}$  do
2:   for  $i = 1$  to  $N$  do
3:      $j := (i + d) \bmod N$ 
4:      $\mathbf{T}_{i \rightarrow j} := \mathbf{P}_j^{-1} \cdot \mathbf{P}_i$ 
5:      $o_{ij} := \text{OVERLAP}(\mathbf{f}_i, \mathbf{f}_j, \mathbf{T}_{i \rightarrow j})$ 
6:     if  $o_{ij} > t_o$  then
7:        $\mathbf{T}_{i \rightarrow j}, e := \text{CLSREGISTRATION}(\mathbf{f}_i, \mathbf{f}_j, \mathbf{T}_{i \rightarrow j}, o_{ij})$ 
8:       if  $e \leq \text{MEDIANRANGE}(\mathbf{f}_i) \cdot t_r + t_a$  then
9:          $\mathbf{G} := \mathbf{G} \cup \{\text{EDGE}(i, j, \mathbf{T}_{i \rightarrow j})\}$ 
10:      end if
11:    end if
12:  end for
13:   $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N = \text{OPTIMIZE}(\mathbf{G})$ 
14: end for
15: return  $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N$ 

```

(in the same way as described above) with additional edges in the pose graph representing the global position in some geodetic frame, as shown in Fig. 2.10b.

2.8 POSE GRAPH VERIFICATION

After the registration is performed, a new edge is added into the pose graph only if the registration error is below a certain threshold modeled by Equation (2.7) (Line 8 of Algorithm 1). However, this simple rejection is not robust enough—some registrations are falsely rejected or accepted. After all overlapping frames are registered, additional verification is performed for all edges.

Expected transformation \mathbf{T}_{ij}^e is computed (Equation (2.9)) using alternative path $\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_{K-1}, \mathbf{T}_K$, as described in Fig. 2.11. The L2 norm of positional difference between expected transformation \mathbf{T}_{ij}^e and the transformation \mathbf{T}_{ij} found by registration (Equations (2.10)–(2.12)) is considered as the error value related to this edge. Note that the positional difference is also affected by the difference in rotation and therefore it is included in this error.

$$\mathbf{T}_{ij}^e = \mathbf{T}_1 \cdot \mathbf{T}_2 \cdot \dots \cdot \mathbf{T}_{K-1} \cdot \mathbf{T}_K \quad (2.9)$$

$$\Delta_{ij} = \mathbf{T}_{ij}^{-1} \cdot \mathbf{T}_{ij}^e \quad (2.10)$$

$$\Delta_{ij} = [\mathbf{R}_{ij} | \mathbf{t}_{ij}] \quad (2.11)$$

$$e_{ij} = \|\mathbf{t}_{ij}\|_2 \quad (2.12)$$

For each edge, all alternative paths up to a certain length are found and their errors are estimated. We use paths of length up to 3 as a tradeoff between the time complexity and robustness. An edge is rejected when the median of these error values is below accepted threshold (10 cm in our experiments). This cannot be considered as

target error of our reconstruction since the pose graph optimization process further decreases the cumulative error. The whole process is repeated until there is no edge to reject.

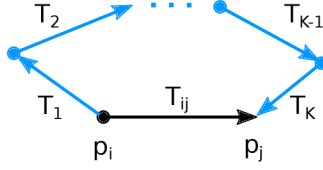


Figure 2.11: Verification of edge (p_i, p_j) representing transformation T_{ij} is performed by comparison with transformation $T_1 \cdot T_2 \dots T_K$ of alternative path (blue) between i th and j th node.

2.9 HORIZONTAL ALIGNMENT OF THE INDOOR MAP

While, for outdoor environment, the model is georeferenced and aligned with NED geodetic coordinate frame (north, east, and down), there is no such possibility when mapping indoors since the GNSS signal is not available. However, practical indoor applications of our 3D mapping solution require at least horizontal alignment—the alignment of gravity vector with Z-axis and the alignment of straight floors/ceilings with XY-plane in resulting 3D model as Fig. 2.12 shows.

This alignment is possible, since roll and pitch angles are provided by IMU (using measurements by accelerometers and gyroscopes) and extrinsic calibration of Velodyne sensors to the IMU frame C_I estimated as described in Sec. 2.3. The simplest solution would be to use these roll and pitch angles directly to align the LiDAR scans individually and deploy the SLAM only to estimate the remaining parameters (heading and translation). Unfortunately, this is not possible because the accuracy of roll and pitch angles is not sufficient—error in order of degrees happens during the motion. Since our goal is to reduce the cost of our solution, we did not want to use additional expensive hardware. We rather propose an alternative approach to estimate horizontal alignment from these noisy measurements.

We can leverage the fact that there are multiple (thousands) of roll/pitch measurements and only a single transformation for horizontal alignment needs to be computed. First, we are able to split each transformation (for each LiDAR frame) estimated by SLAM into the rotation and the translation

$$\mathbf{P}_{\text{SLAM}} = [\mathbf{R}_{\text{SLAM}} | \mathbf{t}_{\text{SLAM}}]. \quad (2.13)$$

Our partial goal is to estimate horizontal alignment \mathbf{A}_h fulfilling Equation (2.14). The transformation of point cloud data X by SLAM rotations \mathbf{R}_{SLAM} and horizontal alignment \mathbf{A}_h is the same, as the transformations of these data by IMU measured rotation \mathbf{R}_{IMU} (including the calibration C_I). In addition, each rotation (SLAM or IMU provided) can be split into roll \mathbf{R}^R , pitch \mathbf{R}^P and heading \mathbf{R}^H (Equation (2.15)).

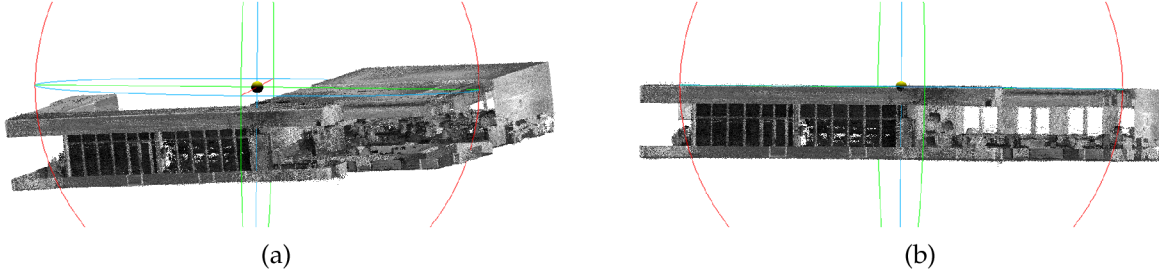


Figure 2.12: The reconstruction built by our SLAM solution before (a) and after (b) the alignment of horizontal planes (floor, ceiling, etc.) with XY plane (blue circle).

Since the IMU sensor is not able to provide accurate heading information indoors, we supplement the heading $\mathbf{R}_{\text{SLAM}}^{\text{H}}$ estimated by SLAM.

$$\mathbf{R}_{\text{IMU}} \cdot \mathbf{C}_I \cdot \mathbf{X} = \mathbf{A}_h \cdot \mathbf{R}_{\text{SLAM}} \cdot \mathbf{X} \quad (2.14)$$

$$\mathbf{R}_{\text{SLAM}}^{\text{H}} \cdot \mathbf{R}_{\text{IMU}}^{\text{P}} \cdot \mathbf{R}_{\text{IMU}}^{\text{R}} \cdot \mathbf{C}_I = \mathbf{A}_h \cdot \mathbf{R}_{\text{SLAM}} \quad (2.15)$$

$$\mathbf{A}_h = \mathbf{R}_{\text{SLAM}}^{\text{H}} \cdot \mathbf{R}_{\text{IMU}}^{\text{P}} \cdot \mathbf{R}_{\text{IMU}}^{\text{R}} \cdot \mathbf{C}_I \cdot \mathbf{R}_{\text{SLAM}}^{-1} \quad (2.16)$$

Using Equation (2.16), we are able to estimate the (noisy and inaccurate) horizontal alignment \mathbf{A}_h for each pair of SLAM and IMU provided rotations of the same timestamp. During the mapping, there are usually thousands of these pairs (10 pairs per second) which are synchronized. The precise horizontal alignment is then computed by averaging the quaternions [9] representing noisy partial alignments \mathbf{A}_h .

2.10 INTENSITIES NORMALIZATION

Another quality we would like to introduce into the 3D model is the approximate surface “color” information to improve the ability of visual recognition of various objects (inventory, signs, etc.). To avoid additional HW, and preserve invariance to illumination conditions, we use the laser return intensity. However, these intensity values cannot be directly considered as surface reflectivity, since they are affected by various additional factors such as angle of incidence, range of the measurement or gain of the particular laser beam. These factors were reported by previous works [5, 7, 6] and also confirmed by our experiments in Fig. 2.13.

Previously published works propose various closed-form solutions of intensity normalization for long range measurements (over 10 m) [5, 7, 6]. However, this is not applicable for smaller indoor environments and therefore we propose an alternative solution. If the normalized intensity represents only the surface reflectivity, there should be no dependency on other factors and probability distribution of the intensities should be the same for different laser beams, angles of incidence, or ranges.

Therefore, we discretize the space of ranges and angles with some small resolution (e.g., 20 cm and 1° , respectively) and we distribute all the points of the point cloud model into a 3D grid based on the source beam ID (already discrete), the angle of incidence and the range. Our goal is to achieve that the intensity probability distribution will be the same for each bin of points. Assuming normal distribution of surface

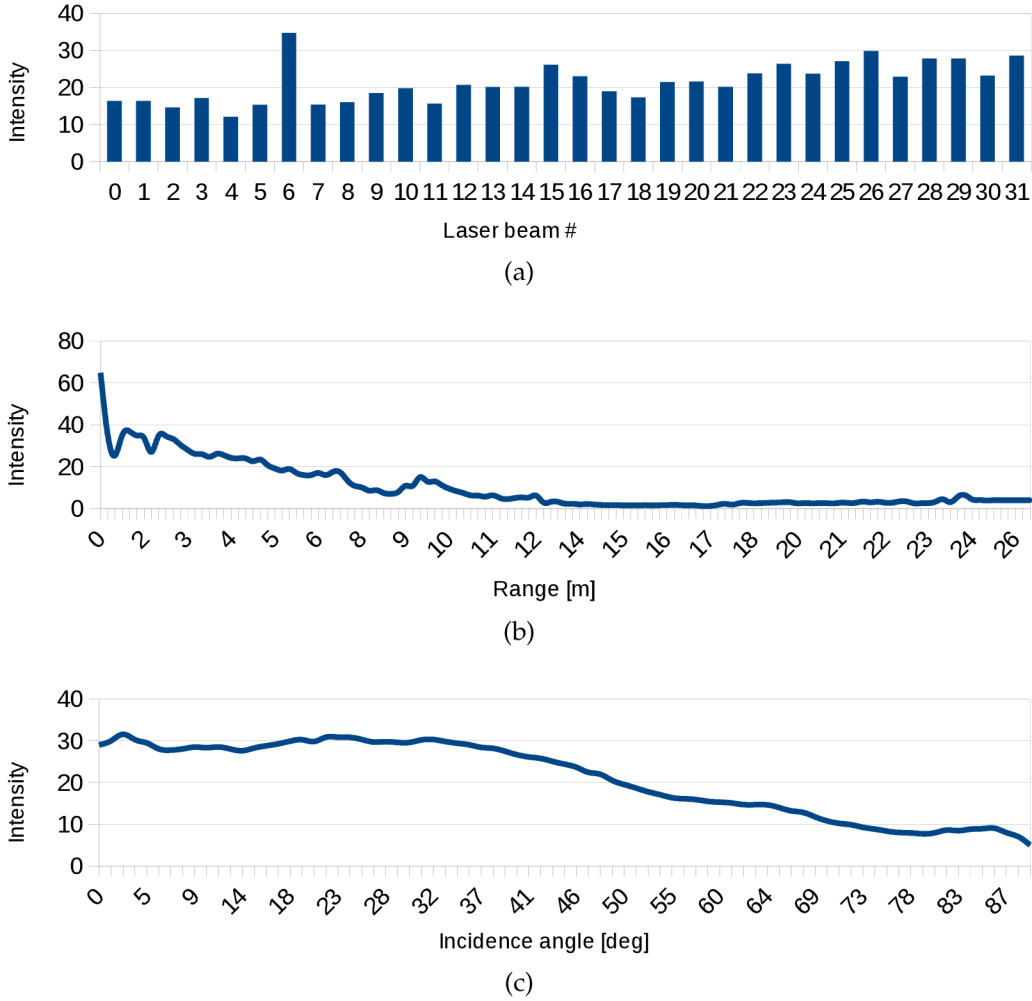


Figure 2.13: The dependency of laser return intensity on: the source beam (a); range of the measurement (b); and the angle of incidence (c). We are using 2 LiDAR scanners with 16 laser beams per each scanner, 32 beams in total.

reflectivities (“colors”), the same target distribution $\mathcal{N}(\mu, \sigma^2)$ will be achieved within each bin by a simple transformation:

$$\mathcal{N}(\mu, \sigma^2) = \mathcal{N}(\mu_i, \sigma_i^2) \cdot \frac{\sigma}{\sigma_i} + (\mu - \mu_i), \quad (2.17)$$

where $\mathcal{N}(\mu_i, \sigma_i^2)$ is the original distribution of laser intensities within i th bin.

There are no ground truth data to perform any objective evaluation of our proposed method for intensity normalization. We are only able to compare the results of 3D reconstruction with and without the normalization. Examples of results can be found in Fig. 2.14.

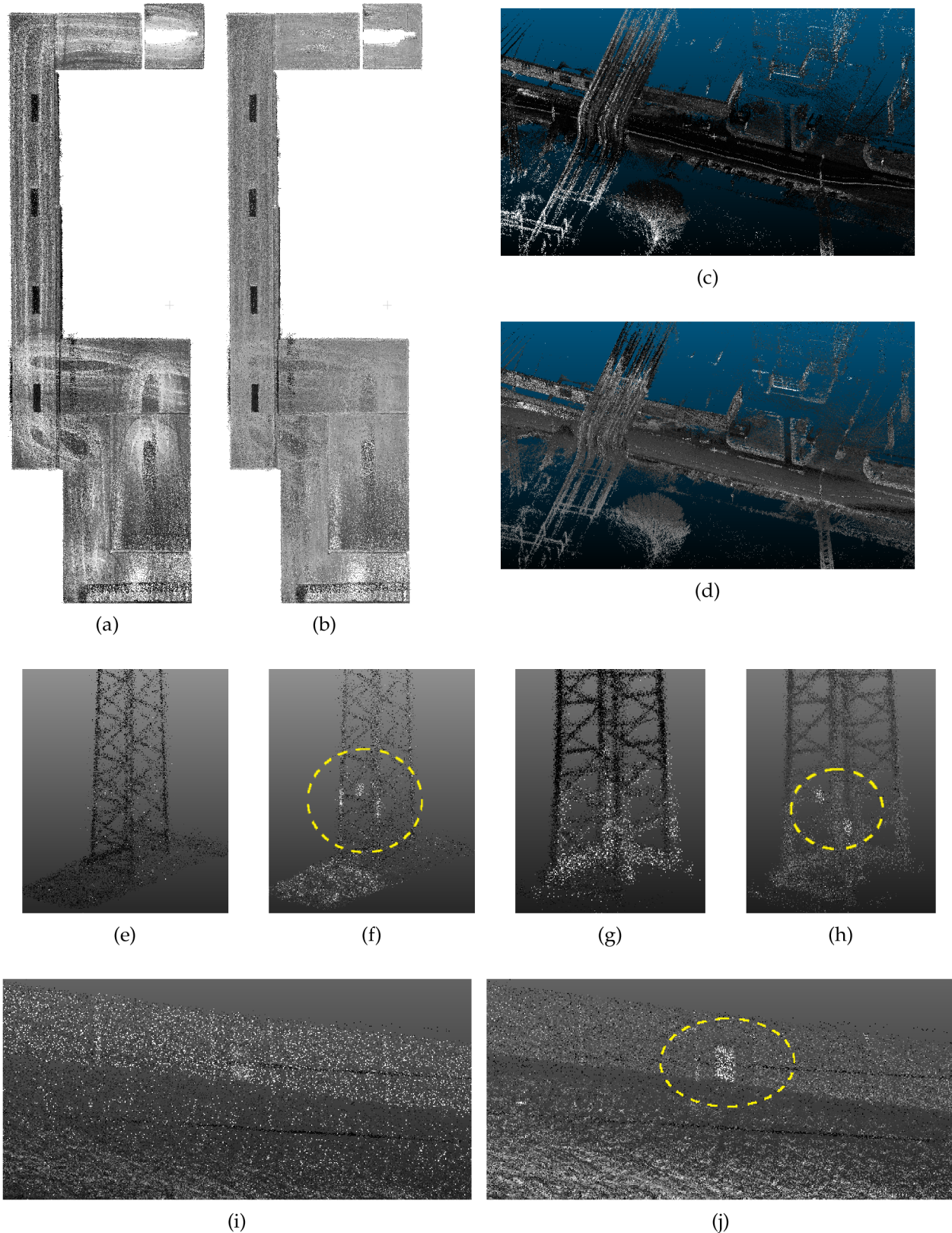


Figure 2.14: Results of 3D reconstruction without (a), (c), (e), (g), (i) and with (b), (d), (f), (h), (j) the normalization of laser intensities. One can observe more consistent intensities for solid color ceiling (b) reducing the artifacts of trajectory, while preserving the contrast with ceiling lights. Besides the consistency, normalization of intensities reduces the noise (d). The most significant improvement is the visibility of important objects e.g., markers at the electrical towers (f), (h) or emergency exit doors (j) at the highway wall. All these objects can't be found in the original point clouds (e), (g), (i).

This section presents mapping results of our system in various scenes and scenarios—outdoor environments where GNSS is available, indoor scenes with GNSS denied, small rooms, staircases, and a narrow corridor. A usable and precise solution must avoid so called “double walls”, which are a typical issue in 3D reconstructions causing ambiguity. Unfortunately, evaluation of such duplicities cannot be performed automatically, thus the operator (a certified geodesist) verified the reconstructions for us by inspecting multiple slices across the model. Moreover, the data density and point coloring by the intensity readings are required for better visual recognition of various objects in the environment. All the raw data collected by our backpack solution, and also the 3D reconstructions used in this evaluation, are publicly available¹.

Regarding the precision, our goal is to achieve 5 cm relative precision (e.g., distance of the point from ground truth) denoted as e_r . For outdoor environments, there are also constraints for absolute error e_a in global geodetic frame. The average of this absolute error is required to be below 14 cm for position in horizontal plane and 12 cm for height estimation. However, the constraints for maximal error are set to double of these values—up to 28 cm for horizontal and 24 cm vertical error. These values were obtained through consultation with experts in the field of geodesy and follow the requirements for creating the building models, outdoor vector maps, inventory check, etc. Global error constraints are applicable only outdoors, where some global positioning system is available. To sum up, in this section, we show that our solution provides:

- sufficient relative precision e_r under 5 cm;
- global absolute error e_a within the limits described above;
- data density and coloring by normalized intensities for visual inspection; and
- data consistency without ambiguity (no dual walls effects).

3.1 COMPARISON OF POINT CLOUD REGISTRATION METHODS

We compared our previously published CLS method [16] with different modes (online and offline) of state-of-the-art method LOAM [17] using the data of KITTI Odometry Suite [1] providing both the Velodyne LiDAR data and ground truth poses. The error metrics used in this evaluation are defined by the KITTI dataset itself. The data sequences are split into subsequences of 100, 200, ..., 800 frames (of 10, 20, ..., 80 s duration). The error e_s of each subsequence is computed as:

$$e_s = \frac{\|\mathbf{E}_s - \mathbf{C}_s\|_2}{l_s}, \quad (3.1)$$

(provided by [1]) where \mathbf{E}_s is the expected position (from the ground truth) and \mathbf{C}_s is the estimated position of the LiDAR where the last frame of subsequence was

¹ <http://www.fit.vutbr.cz/~ivelas/files/4RECON-dataset.zip>

		Error e_s (3.1)			
Sequence	Length	LOAM Online	LOAM Offline	CLS Single	CLS Multi-Frame
0	4540	0.052	0.022	0.022	0.018
1	1100	0.038	0.040	0.042	0.029
2	4660	0.055	0.046	0.024	0.022
3	800	0.029	0.019	0.018	0.015
4	270	0.015	0.015	0.017	0.017
5	2760	0.025	0.018	0.017	0.012
6	1100	0.033	0.016	0.009	0.008
7	1100	0.038	0.019	0.011	0.007
8	4070	0.035	0.024	0.020	0.015
9	1590	0.043	0.032	0.020	0.018
Weighted average	2108	0.043	0.029	0.022	0.017

Table 3.1: Comparison of visual odometry error for SoA method LOAM and our CLS method. The experiments were performed on KITTI Odometry dataset [1]. For CLS, frame to frame (single) or frame to multiple (10) neighboring frames (multi-frame) registrations without any loop closures were performed. In LOAM experiments, both the original online version (providing real time performance) and offline version (with full procedure for each frame omitting approximations) was used. In all data sequences, except the short sequence No. 4 where the car drives only forward without any turns, our multi frame approach outperformed the LOAM solution.

taken with respect to the initial position (within given subsequence). The difference is divided by the length l_s of the followed trajectory. The final error value is the average of errors e_s across all the subsequences of all the lengths.

The experiment is summarized in Table 3.1 and it leads to the conclusion that our CLS approach outperforms LOAM with approximately 1 cm lower drift per 1 m of trajectory elapsed. For clarification, LOAM can run in two different modes. In the online mode (10 fps), mapping is skipped for a certain number of frames, which are only roughly aligned. In the offline mode, which is approximately $3\times$ slower, every frame undergoes the full mapping procedure.

The precision of our method was estimated for frame-to-frame approach, where only consequent frames were registered, and also for the scenario, where each frame is registered with all other frames within a small neighborhood (10 neighboring frames used in this experiment). In this experimental multi-frame approach, the final pose is estimated by simple averaging.

In our previous publication [16], the superior performance of CLS over GICP method (Generalized ICP) [12] was presented, too. All these evaluations led to the choice of CLS for the LiDAR frames registration in our 4RECON backpack solution.

3.2 INDOOR EXPERIMENTS

For indoor evaluation of our system, we chose two different environments—the office and staircase in Fig. 3.1—where our partner company has already performed 3D

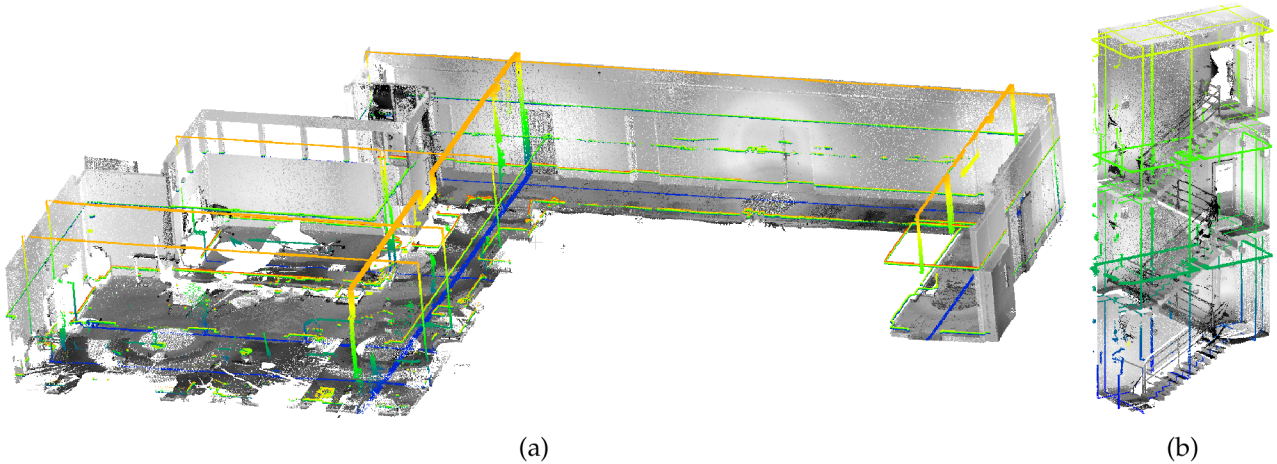


Figure 3.1: Experimental environments Office (a) and Staircase (b), and the highlighted slices that were used for precision evaluation.

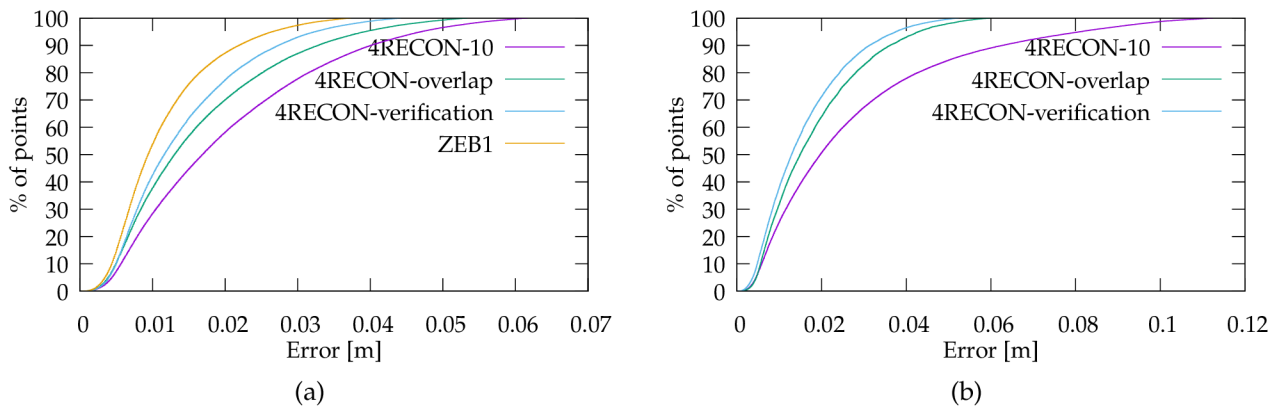


Figure 3.2: Error e_r distribution (the amount of the points within certain error) for our system $4RECON$ and $ZEB-1$ product. The experiments were performed for all test slices in Fig. 3.1 on Office (a) and Staircase (b) dataset. Note that the model built by $ZEB-1$ was not available and therefore the evaluation is missing.

mapping using different laser scanners and generously provided the accurate output models to us. The reconstructions from static *FARO* scanner achieving very high accuracy (in order of millimeters) were used as the ground truth. The same strategy has been already used for evaluation of other mapping systems [10, 14, 8]. For the office environment only, also the 3D reconstruction created by *ZEB-1* solution was provided to us. This allowed us to compare our solution in terms of accuracy, data density, model usability and completeness.

To evaluate the relative error, all the models of the same environment provided by different scanners (*FARO*, *ZEB-1*, and our solution $4RECON$) were aligned using ICP. As displayed in Fig. 3.1, several reference slices (8 slices per model, 16 slices in total) were created for the evaluation of precision. Within each slice, the average error (in Table 3.2) was estimated as the average distance of the 3D points to the ground truth model created by the *FARO* scanner. Our solution achieved approximately 1.5 cm relative error on average, which is only slightly worse result than 1.1 cm error for *ZEB-*

Dataset	Slice #	4RECON-10	4RECON-Overlap	4RECON-Verification	ZEB-1
Office	1	2.50	1.71	1.49	1.44
	2	1.97	1.47	1.31	1.06
	3	1.70	1.75	1.55	1.22
	4	1.82	1.54	1.31	1.22
	5	1.93	1.63	1.53	1.44
	6	2.13	1.49	1.47	1.29
	7	2.09	1.68	1.37	0.97
	8	2.07	1.36	1.37	1.31
Average e_r (cm)		2.01	1.62	1.41	1.14
Staircase	1	3.23	2.11	1.81	-
	2	3.99	1.87	1.60	-
	3	2.63	1.65	1.61	-
	4	2.74	1.71	1.53	-
	5	2.42	1.68	1.50	-
	6	2.98	2.67	1.67	-
	7	1.76	1.75	1.29	-
	8	1.82	1.67	1.56	-
Average e_r (cm)		2.74	1.82	1.57	-

Table 3.2: Relative error e_r of our method and ZEB-1 product within selected slices visualized in Fig. 3.1. Presented values are average displacements (cm) of the points comparing with the ground truth point cloud obtained by FARO static scanner. The results are missing for ZEB-1 and Staircase dataset since there was no reconstruction using this scanner available.

1 that is burdened by the multiple limitations listed below in this section. Moreover, we provide information about the distribution of displacement relative error in Fig. 3.2. The error was estimated for ZEB-1 and different modes of our system:

- in *4RECON-10*, the registrations were performed only within small neighborhood of 10 nearest frames (1 s time window) and reflects the impact of accumulation error;
- for *4RECON-overlap*, the registrations were performed for all overlapping frames as described in Sec. 2.7 reducing the accumulation error by loop closures at every possible location; and
- pose graph verification (see Sec. 2.8) was deployed in *4RECON-verification*, yielding the best results with good precision and no ambiguities.

Both ZEB-1 and our solution including pose graph verification achieved sufficient accuracy below 5 cm. Moreover, the precision of 2 cm was fulfilled for more than 70% of data. Slightly better precision of ZEB-1 solution was achieved thanks to the Hokuyo sensor with $4\times$ higher scanning frequency while preserving much lower vibrations compared with Velodyne LiDAR.

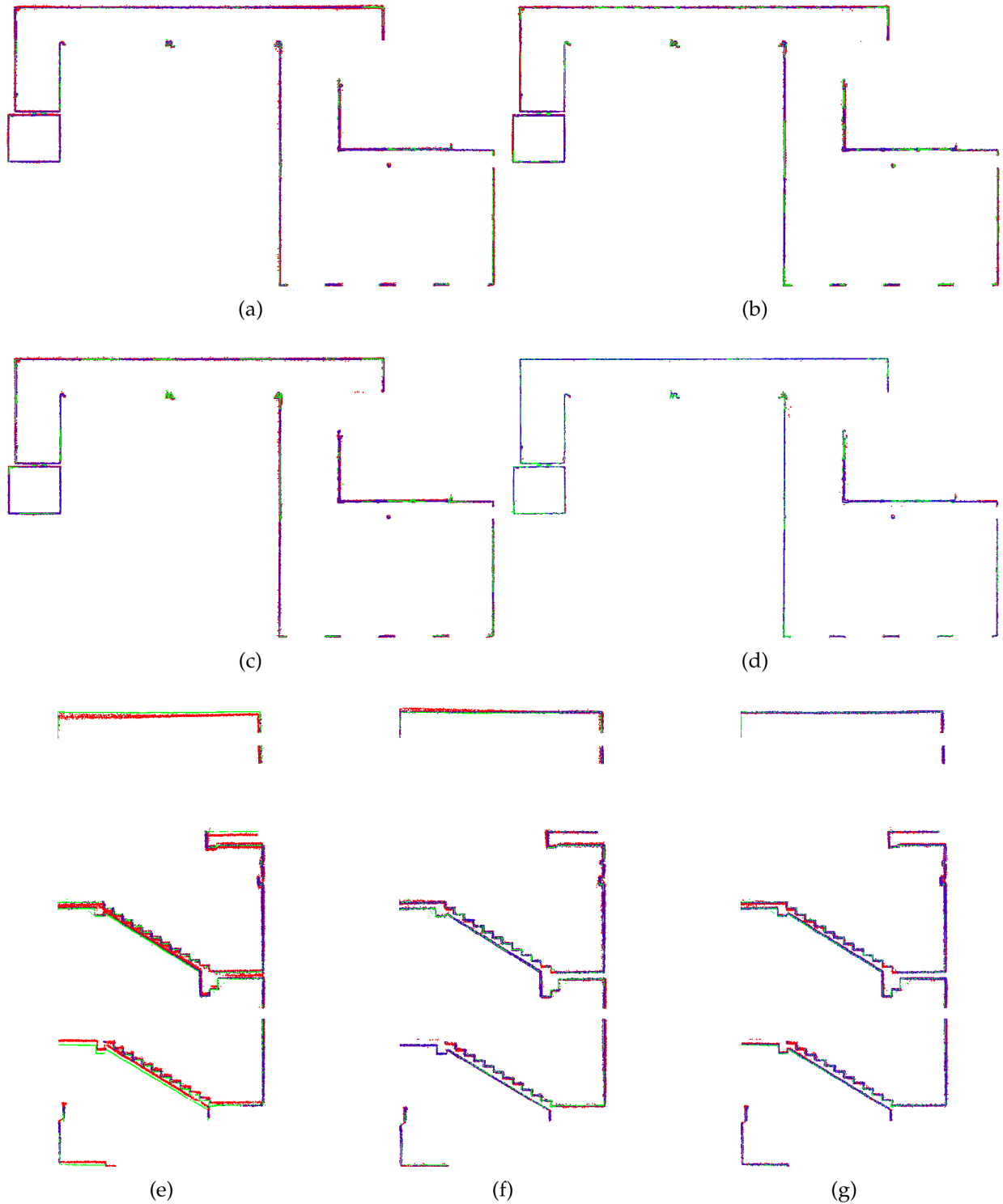


Figure 3.3: Color coded errors within the horizontal reference slice of the Office dataset (a)–(d) and vertical slice in Staircase dataset (e)–(g). Blue color represents zero error, red color stands for 10 cm error and higher. The ground truth FARO data are displayed in green. The results are provided for 4RECON-10 (a,e), 4RECON-overlap (b,f), 4RECON-verification (c,g), and ZEB-1 (d). For Office dataset, there are no ambiguities (double walls) even without visual loop detection while both loop closure and pose graph verification is necessary for more challenging Staircase dataset to discard such errors. Moreover, one can observe that ZEB-1 solution yields lower noise reconstruction thanks to the less noisy Hokuyo LiDAR.

Fig. 3.3 also shows the precision within representative slices—horizontal slice for Office dataset and vertical slice across model of Staircase. These slices demonstrate the noise within data coming from different sensors—Hokuyo LiDAR for ZEB-1 solution and Velodyne for our 4RECON system—and also the precision for different modes of operation. For Staircase dataset, the necessity of pose graph optimization is also demonstrated.

Our evaluations show that the precision of our 4RECON backpack is comparable to the solution ZEB-1 while fulfilling basic requirement for relative error below 5 cm. Note that the error values are also comparable (and in some cases better) to the precisions of other available solutions. In our solution, higher noise can be observed comparing with ZEB-1. This corresponds with higher error values and it is the main reason for little lower accuracies.

However, it is important to point out two most significant advantages of our solution comparing with ZEB solutions. First, our solution is *usable in vast open spaces* with fewer and more distant featuring objects, as is demonstrated in the next sections. In indoor environments featuring objects at distances significantly larger than 15–20 m [2], ZEB solutions based on the Hokuyo sensor fail.

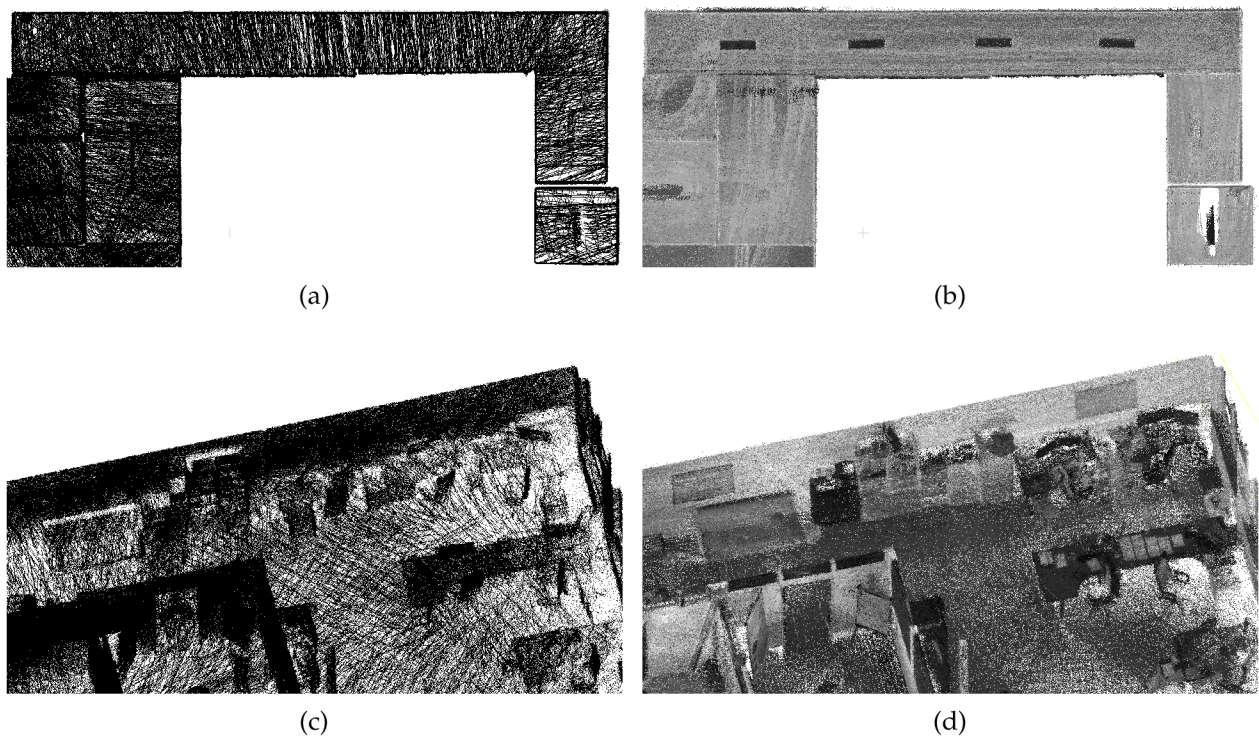


Figure 3.4: The comparison of data density provided by ZEB-1 (a,c) and our (b,d) solution. Since the ZEB-1 solution is based on the Hokuyo scanner, the laser intensity readings are missing and data density is much lower compared with our solution. Multiple objects which can be distinguished in our reconstruction (lamps on the ceiling in the top, furniture and other equipment in the bottom image) are not visible in the ZEB-1 model.

Second, our Velodyne-based solution is able to provide much higher *data density*, *map completeness* and *visibility of objects* in the scene. We chose two large surfaces (the ceiling and the side wall in Fig. 3.4) with 230 m^2 in total area. Models of these surfaces created by ZEB-1 solution achieved average data density 0.9 points per cm^2

(2.2 million points in total). Models created by our 4RECON backpack consist of more than 23 million points, achieving much higher data density—10.1 points per cm^2 . Better visibility of objects in Fig. 3.4 is achieved thanks to the laser intensity readings provided by Velodyne sensor and employing our normalization process as described in the Sec. 2.10. This might appear to be only a “cosmetic” property, but the visibility of the construction elements, equipment, furniture, etc. in the scene is important for usability in real applications—e.g., an operator needs to distinguish between the window and the blackboard.

3.3 OUTDOOR EXPERIMENTS

Our system is a universal solution—both for indoor scenes, where the usability was proven by the previous section, and for outdoor scenes, including vast open ones. We tested and evaluated our system during a real task—high voltage lines mapping and measurement. The area of interest, including the details of some important objects, is visualized in Fig. 3.5. The main goal of this mission was position estimation of electric pylons (including footprint of the base, total height and the positions of the wire grips) and the heights and the hangings of the wires. Fig. 3.5 shows that these details can be recognized in the 3D model. The usability of our 3D reconstructions was also confirmed by the geodetic company we asked for manual data inspection and evaluation.

In the same way as during the indoor mapping, the ambiguities in multiple instances of objects disqualifies the reconstructions to be used in practical geodetic measurements. Such error in comparison with the desired result of the reconstruction is shown in Fig. 3.6. Multiple instances of the same object, blurred and noisy results were successfully avoided by our solution (see Figures 3.5 and 3.6).

Since our solution integrates precise GNSS/INS module for outdoor scenarios, the model is *georeferenced*—the coordinates of all the points are bound in some global geodetic frame.

To verify the absolute positional accuracy of our model, we performed precise measurements on so-called survey markers. This is commonly used technique to verify the precision of resulting maps (including 3D maps). Precise positions of the survey markers are estimated using specialized geodetic GNSS system, which is placed statically on the survey point for several seconds, until the position converged. The precision up to 2 cm is achieved using RTK (Real Time Kinematics) which are received online via internet connection.

Survey markers (Fig. 3.7a) are highlighted using high-reflective sprays. Thanks to the coloring of point cloud by laser intensities, these markers are also visible in the reconstructions as can be seen in Fig. 3.7b.

The evaluation in Table 3.3 shows that our 3D mapping for 0.5 km test track fulfills the requirements for absolute error, as described at the beginning of this section—average error below 14 cm for position in horizontal plane and 12 cm for height estimation and maximal error up to 28 cm and 24 cm, respectively (double values of expected average error).

Thanks to the ability of point cloud coloring by laser intensities, it is possible to also run such evaluation for the validation of each 3D model, which should be used

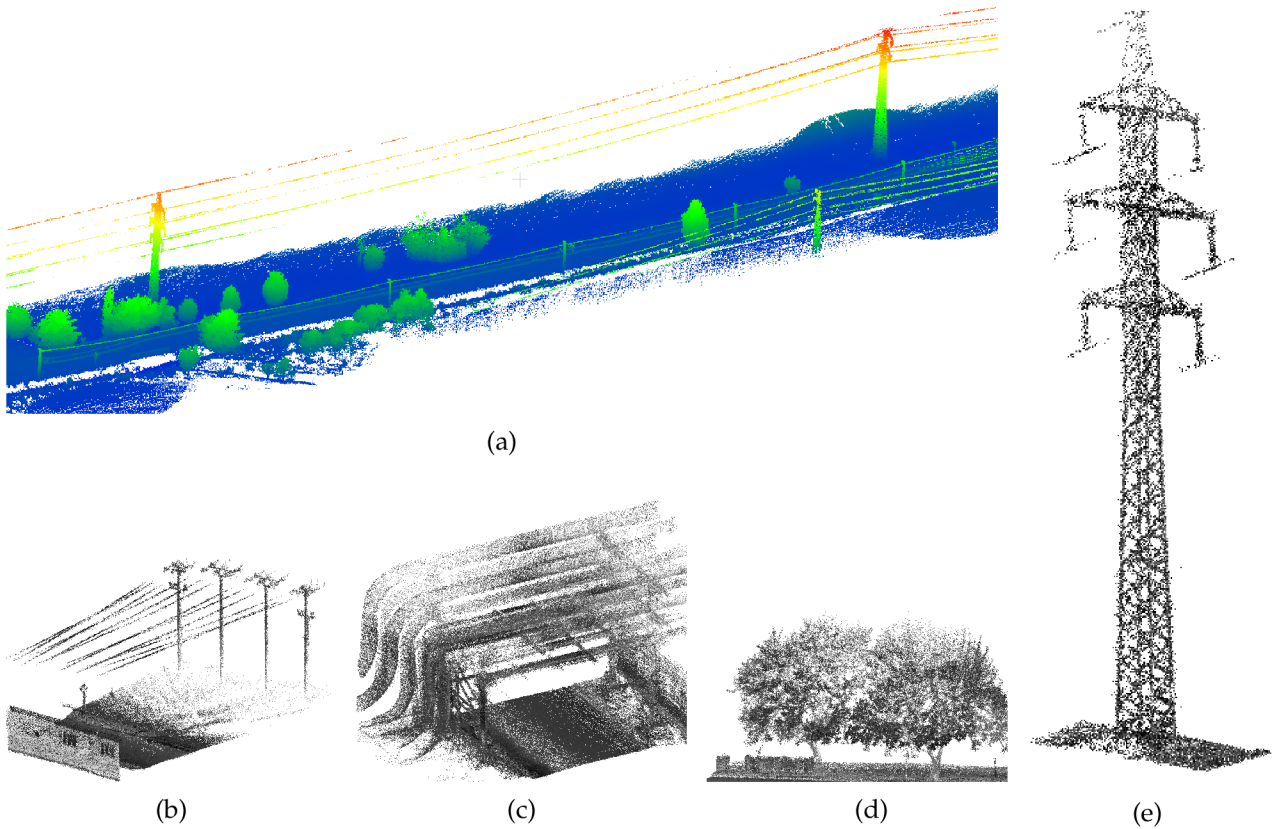


Figure 3.5: The example of 3D reconstruction of open field with high voltage electrical lines (a). The model is height-colored for better visibility. The estimation of positions and height of the lines (b), towers (e), etc. was the main goal of this mapping task. The other elements (c,d) in the scene are shown for demonstration of the reconstruction quality.

in real application. This is also an important quality, since there are requirements for double measurements in geodesy to ensure that the accuracy is sufficient.

Table 3.3: Errors measured (cm) on geodetic survey marker points at the beginning and at the end of survey track. The distance between the control points is 523 m.

Ref. Point	dX	dY	Horizontal Error	dZ (Vertical)	Total Error e_a
1	-5.9	-1.2	6.0	-15.2	16.3
2	-5.6	0.5	5.6	-4.7	7.3

3.4 COMPARISON OF SINGLE AND DUAL VELODYNE SOLUTION

Finally, we compared the robustness of our dual LiDAR solution over the system with single LiDAR only. We computed reconstructions of the Office environment using our solution with two synchronized and calibrated LiDARs (one aligned vertically and second horizontally) in Figure 3.8a,b and also using only single LiDAR—horizontally (Figure 3.8c,d) or vertically aligned (Figure 3.8e,f).

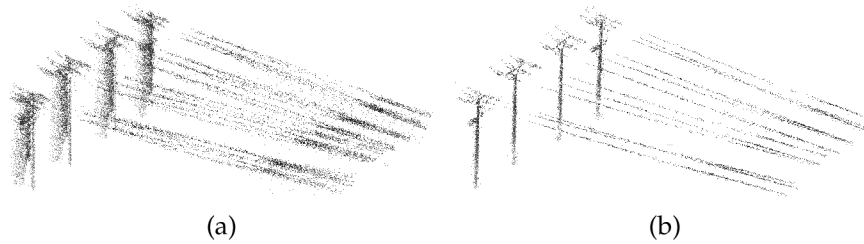


Figure 3.6: Example of ambiguities caused by reconstruction errors (a), which disqualifies the model to be used for practical measurements. We obtained such results when we used only poses provided by GNSS/INS subsystem without any refinements by SLAM or point cloud registration. Our solution (including SLAM) provides valid reconstructions (b), where both towers and wires (in this case) can be distinguished.

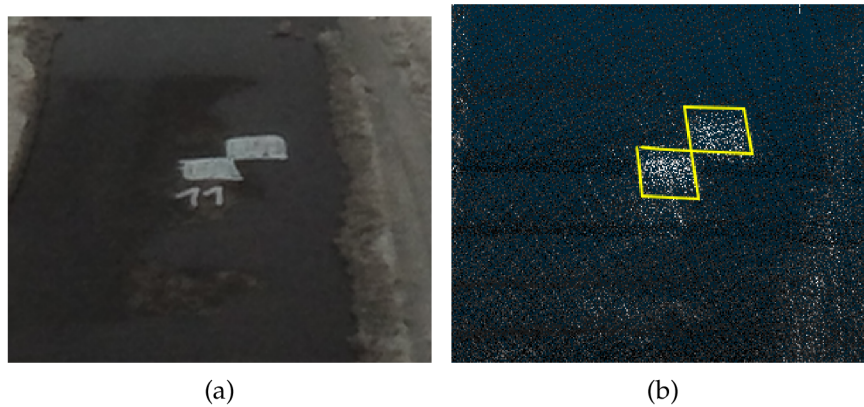


Figure 3.7: Geodetic survey markers painted on the road (a) is also visible in the point cloud (b) thanks to the coloring by laser intensities.

Our evaluation shows that the dual LiDAR solution provides a valid reconstruction. However, the solution with horizontal LiDAR only is not able to provide vertically correct alignment (Figure 3.8d), and vice versa, the solution with vertical LiDAR is horizontally misaligned (Figure 3.8e).

3.5 DISCUSSION

When we look on the 4RECON mapping backpack in the context of the other available solutions, we can summarize its advantages and disadvantages.

Comparing to the ZEB products, our backpack achieves much higher data density, better visibility of the objects in the resulting model, higher comfort of data acquisition, and, most importantly, usability also in the outdoor featureless open spaces, including the option of georeferencing the reconstructed point map. However, we must admit that ZEB scanners achieve better accuracy and lower noise in the models of indoor environments.

In terms of universality of the usage, our solution also outperforms Robin and Akhka backpacks, which require GNSS readings and therefore indoor scanning is not

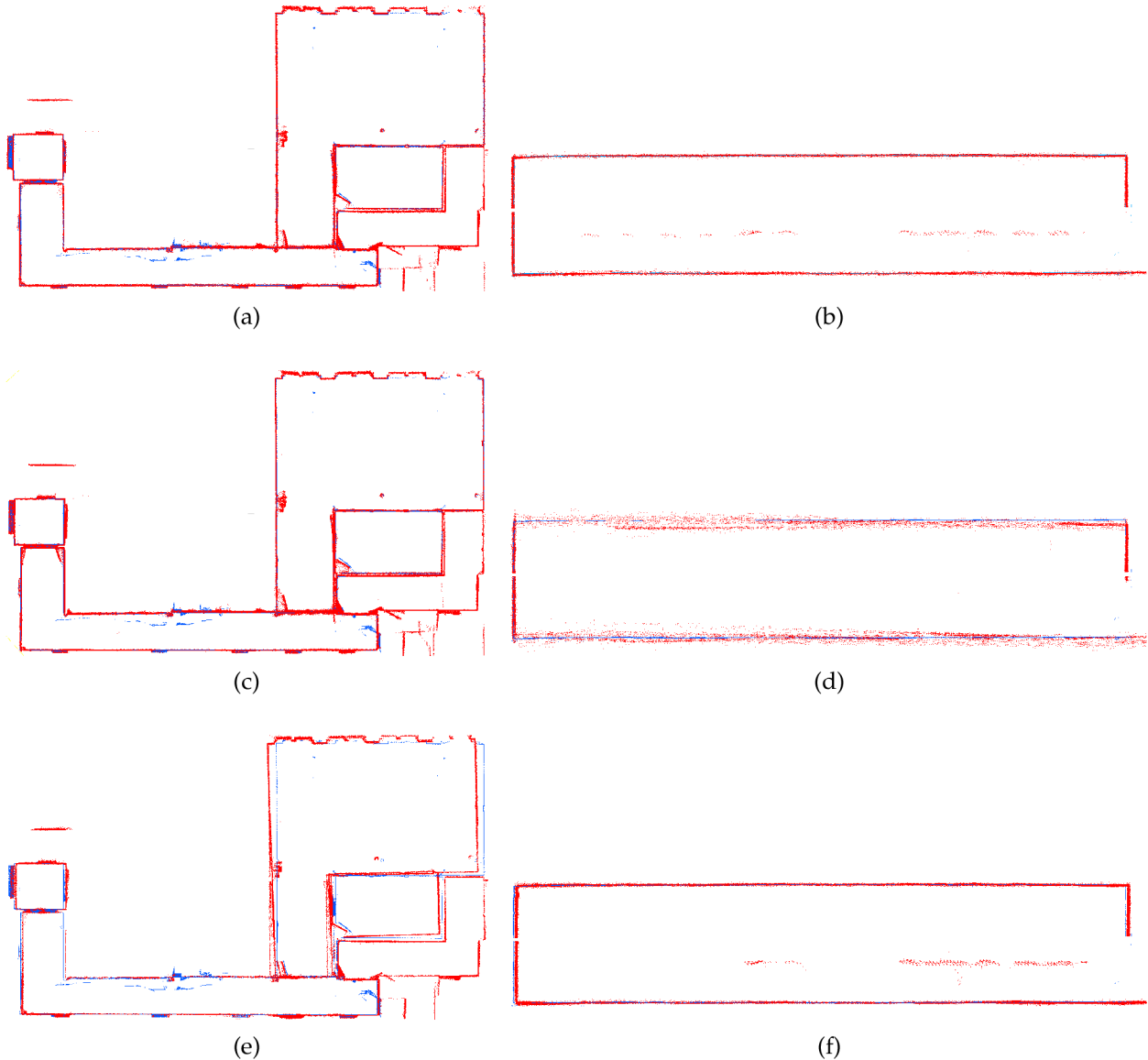


Figure 3.8: Comparison of reconstructions provided by dual LiDAR system—floor plan top view (a) and side view of the corridor (b)—with the reconstruction built using only single horizontally (c,d) or vertically (e,f) positioned Velodyne LiDAR. The reconstructions are red colored with ground truth displayed in blue.

possible. For outdoor tasks, Robin achieves better precision than our 4RECON backpack, but it is also important to point out the very high price of the Robin solution.

Laser mapping backpacks Pegasus, Viametris bMS3D and LiBackpack can be considered as the most similar solutions to our work. All these systems claim precision up to 5 cm, which is also the accuracy of 4RECON (according to the evaluation in Fig. 3.1). The advantages of these solutions are more professional design and the presence of additional RGB cameras (for Pegasus and Viametris backpacks). The integration of panoramic RGB camera into our backpack is the plan for future work. Our solution on the other side provides open SLAM method in comparison with the proprietary solutions deployed in these backpacks, and also potentially much lower price.

CONCLUSIONS

This work presents a dual LiDAR system for mobile mapping. Our solution can be easily carried as a backpack together with a reliable dual antenna GNSS/INS system. This leads to the universality of its usage. In small or narrow indoor environments with many obstacles, two LiDAR sensors increase the field of view. On the other side, in open outdoor spaces with lack of features, the reliable positional subsystem keeps the result accurate.

Thanks to the type of LiDARs used, our solution also brings multiple other beneficial properties: data density, map completeness and coloring by laser intensities normalized by our novel algorithm. The intensities enables better visual recognition of the elements in the scene as well as the visibility of geodetic survey markers for checking the model validation.

The proposed solution was evaluated in both indoor and outdoor scenarios. During the mapping of the office or staircase environment, our solution fulfilled the requirement of error below 5 cm and achieved a similar precision as solution ZEB-1. The average error in terms of the points displacements is approximately 1.5 cm. For outdoor experiments, our reconstruction met the requirements for absolute precision with 11.8 cm average error in the global geodetic frame. This proves higher universality of our mapping backpack compared to the previous ZEB-1 solution. In all our experiments, data consistency was preserved and unambiguous models were built.

BIBLIOGRAPHY

- [1] A Geiger, P Lenz, C Stiller, and R Urtasun. Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013. doi: 10.1177/0278364913491297.
- [2] *The ZEB-REVO Brochure*. GeoSLAM Ltd., apr 2018. Available at https://geoslam.com/wp-content/uploads/2018/04/GeoSLAM-ZEB-REVO-Solution_v9.pdf?x97867.
- [3] *The ZEB-REVO Solution*. GeoSLAM Ltd., 2018. Available at <http://download.geoslam.com/docs/zeb-revo-rt/ZEB-REVO%20RT%20User%20Guide%20V1-0-1.pdf>, version 9.
- [4] Viorela Ila, Lukas Polok, Marek Solony, and Pavel Svoboda. SLAM++ – a highly efficient and temporally scalable incremental SLAM framework. *The International Journal of Robotics Research*, 36(2):210–230, 2017. doi: 10.1177/0278364917691110.
- [5] Boris Jutzi and H Gross. Normalization of LiDAR intensity data based on range and surface incidence angle. *ISPRS Journal of Photogrammetry and Remote Sensing*, 38, 01 2009.
- [6] Sanna Kaasalainen, Anttoni Jaakkola, Mikko Kaasalainen, Anssi Krooks, and Antero Kukko. Analysis of incidence angle and distance effects on terrestrial laser scanner intensity: Search for correction methods. *Remote Sensing*, 3(10):2207–2221, 2011. ISSN 2072-4292. doi: 10.3390/rs3102207.
- [7] Alireza G. Kashani, Michael J. Olsen, Christopher E. Parrish, and Nicholas Wilson. A review of LiDAR radiometric processing: From ad hoc intensity correction to rigorous radiometric calibration. *Sensors*, 15(11):28099–28128, 2015. ISSN 1424-8220. doi: 10.3390/s151128099.
- [8] Mehdi Maboudi, Dávid Bánhidi, and Markus Gerke. Evaluation of indoor mobile mapping systems. In *GFaI Workshop 3D-NordOst 2017 (20th Applcation-oriented Workshop on Measuring, Modeling, Processing and Analysis of 3D Data)*, 12 2017.
- [9] F. Landis Markley, Yang Cheng, John Lucas Crassidis, and Yaakov Oshman. Averaging quaternions. *Journal of Guidance, Control, and Dynamics*, 30(4):1193–1197, Jul 2007. ISSN 0731-5090. doi: 10.2514/1.28949.
- [10] A Masiero, F Fissore, A Guarnieri, M Piragnolo, and A Vettore. Comparison of low cost photogrammetric survey with TLS and leica pegasus backpack 3D modelss. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42: 147, 2017. doi: 10.5194/isprs-archives-XLII-2-W8-147-2017.
- [11] A. Nuchter, K. Lingemann, J. Hertzberg, and H. Surmann. 6D SLAM with approximate data association. In *Advanced Robotics, 2005. ICAR. Proceedings., 12th Int. Conference on*, pages 242–249, July 2005. doi: 10.1109/ICAR.2005.1507419.
- [12] Aleksandr Segal, Dirk Hähnel, and Sebastian Thrun. Generalized-ICP. In Jeff Trinkle, Yoky Matsuoka, and José A. Castellanos, editors, *Robotics: Science and Systems*. The MIT Press, 2009. ISBN 978-0-262-51463-7. doi: 10.15607/RSS.2009.V.021.
- [13] Ken Shoemake. Animating rotation with quaternion curves. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '85*, pages 245–254, New York, NY, USA, 1985. ACM. ISBN 0-89791-166-0. doi: 10.1145/325334.325242.

- [14] Beril Sirmacek, Yueqian Shen, Roderik Lindenbergh, Sisi Zlatanova, and Abdoulaye Diakite. Comparison of ZEB-1 and leica C10 indoor laser scanning point clouds. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3:143, 2016. doi: 10.5194/isprs-annals-III-1-143-2016.
- [15] Martin Velas, Thomas Faulhammer, Michal Spanel, Michael Zillich, and Markus Vincze. Improving multi-view object recognition by detecting changes in point clouds. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–7, Dec 2016. doi: 10.1109/SSCI.2016.7850045.
- [16] Martin Velas, Michal Spanel, and Adam Herout. Collar line segments for fast odometry estimation from velodyne point clouds. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4486–4495, May 2016. ISBN 978-1-4673-8026-3. doi: 10.1109/ICRA.2016.7487648.
- [17] Ji Zhang and Sanjiv Singh. LOAM: Lidar odometry and mapping in real-time. In *Robotics: Science and Systems Conference (RSS 2014)*, 2014. ISBN 978-0-9923747-0-9. doi: 10.15607/RSS.2014.X.007.

Martin Velás

PhD Candidate

I am focusing on SLAM solutions for mobile laser mapping.



Personal information

Address:

Božetěchova ½, Brno, 61266

Czech Republic

ivelas@fit.vutbr.cz

Tel.: +420775073473

Date/place of birth

15.6.1989

Bánovce nad Bebravou, Slovak Republic

Language skills

- English (B2)
- German (A1)

Other information

Nationality: Slovak

Hobbies: Swimming, cross-country skiing, reading, TV series, and singing.

Education

September 2011 – June 2013

Master degree at Brno University of Technology, Brno, Czech Republic
Faculty of Information Technology, programme: Intelligent systems

September 2008 – June 2011

Bachelor degree at Brno University of Technology, Brno, Czech Republic
Faculty of Information Technology, programme: Information Technology

Work experience

September 2013 - present

PhD Candidate at Brno University of Technology, Brno, Czech Republic
Topic: 3D Mapping from Sparse LiDAR Data

August 2015 - November 2015

Internship at Vienna University of Technology, Vision4Robotics group
Topic: Improvements of Change Detection in Point Cloud Data and
Visualization of the Changes by Incorporating Semantic Information

August 2013 - July 2015

Quality assurance of JBoss Enterprise Application Platform
Red Hat Inc., Brno, Czech Republic

Publications

VELAS Martin, ŠPANĚL Michal, SLEZIAK Tomáš, HABROVEC Jiří and HEROUT Adam. **Indoor and Outdoor Backpack Mapping with Calibrated Pair of Velodyne LiDARs**. *In: Sensors*. 2019, ISSN 1424-8220.

VELAS Martin, ŠPANĚL Michal, HRADIŠ Michal and HEROUT Adam. **CNN for IMU Assisted Odometry Estimation using Velodyne LiDAR**. *In: Proceedings of IEEE International Conference on Autonomous Robot Systems and Competitions*. Torres Vedras: Institute of Electrical and Electronics Engineers, 2018, pp. 71-77. ISBN 978-1-5386-5221-3.

VELAS Martin, ŠPANĚL Michal, HRADIŠ Michal and HEROUT Adam. **CNN for Very Fast Ground Segmentation in Velodyne LiDAR Data**. *In: Proceedings of IEEE International Conference on Autonomous Robot Systems and Competitions*. Torres Vedras: Institute of Electrical and Electronics Engineers, 2018, pp. 97-103. ISBN 978-1-5386-5221-3.

VELAS Martin, ŠPANĚL Michal and HEROUT Adam. **Collar Line Segments for Fast Odometry Estimation from Velodyne Point Clouds**. *In: Proceedings of*

IEEE International Conference on Robotics and Automation. Stockholm: IEEE Computer Society, 2016, pp. 4486-4491. ISBN 978-1-4673-8025-6.

VELAS Martin and ŠPANĚL Michal. **Improving Multi-view Object Recognition by Detecting Changes in Point Clouds**. In: *IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing*. Atény: IEEE Computer Society, 2016, pp. 1-7. ISBN 978-1-5090-4239-5.

VELAS Martin, ŠPANĚL Michal, MATERNA Zdeněk and HEROUT Adam. **Calibration of RGB Camera With Velodyne LiDAR**. In: *WSCG 2014 Communication Papers Proceedings*. Plzeň: Union Agency, 2014, pp. 135-144. ISBN 978-80-86943-71-8. ISSN 1213-6972.

Research projects

April 2019 – Present

Improvements of the precision for mobile 3D mapping system
In cooperation with GEODROM company

October 2018 – February 2019

Experimental Verification of Portable 3D Mapping System Features
In cooperation with GEODROM company

January 2017 – December 2018

Geodrom - **Analysis of Software Tools for 3D Data Processing**
In cooperation with GEODROM company

2017 – 2019

Processing, recognition and imaging of multimedia and 3D data
BUT, FIT-S-17-3984

January 2017 – September 2018

Transfer of knowledge in the field of 3D reconstruction and mapping
Ministry of Industry and Trade of Czech Republic

August 2015 – December 2015

Improvements of Change Detection in Point Cloud Data and Visualization of the Changes by Incorporating Semantic Information
Ministry of Education, Youth and Sports of CR, AKTION, 2015

2014 – 2017

**Algorithms, Design Methods, and Many-Core Execution Platform for
Low-Power Massive Data-Rate Video and Image Processing**

Artemis JU, 7H14002, 621439

2011 – 2015

The IT4Innovations Centre of Excellence

MŠMT ČR, ED1.1.00/02.0070