Localization of Mobile Laser Scanner using Classical Mechanics

Ville V. Lehtola[1], Juho-Pekka Virtanen[1], Antero Kukko[2], Harri Kaartinen[2], Hannu Hyyppä[1]

[1]*Institute of Measuring and Modeling for Built Environment, Aalto University, P.O.Box 15800 , 00076 AALTO, Finland*

[2]*Finnish Geodetic Institute, P.O. Box 15, 02431 Masala, Finland*

Abstract

We use a single 2D laser scanner to 3D scan indoor environments, without any inertial measurement units or reference coordinates. The localization is done directly from the point cloud in an *intrinsic* manner compared to other state-of-the-art mobile laser scanning methods where external inertial or odometry sensors are employed and synchronized with the laser scanner. Our approach is based on treating the scanner as a holonomic system. A novel type of scanner platform, called VILMA, is designed and built to demonstrate the functionality of the presented approach. Results from flat-floor and non-flat-floor environments are presented. They suggest that intrinsic localization may be generalized for broader use.

Keywords: Mobile, Scanner, Platforms, Design, LIDAR, IMU, Localization, Trajectory

1. Introduction

Mobile laser scanning (MLS) answers to challenges where air-borne laser scanners do not reach and terrestrial laser scanning (TLS) is too cumbersome. These challenges include indoor environments (Xiao and Furukawa, 2012; Liu et al., 2010; Holenstein et al., 2011; Thomson et al., 2013; Bosse et al., 2012), but also accurate measurement and modeling of some outdoor environments (Kukko et al., 2012; Bosse and Zlot, 2009; Barber et al., 2008; Kukko, 2013). The Achilles' heel of MLS is localization, or more specifically, the localization of data coming from a moving scanner with six degrees of

freedom. An off-the-shelf answer is to deploy an inertial measurement unit (IMU) to track the movement of the MLS platform, i.e. to obtain its trajectory $j(t)$. This, however, leads into another problem, which is that finite inertial errors accumulate with time and without boundaries. To counter this, outdoor MLS can employ the global navigation satellite system (GNSS) for positioning to provide accurate reference points to the platform trajectory (Kukko et al., 2012). Indoors, however, the GNSS is not available.

In order to cope with the lack of inertial reference frame, such as the GNSS, indoor approaches then require prior knowledge to localize scan data. This prior knowledge typically includes assumptions about the environment that introduce side effects. The flat-floor assumption is commonly used, see e.g. Thomson et al. (2013), although it has problems with, for example, stairs. Xiao and Furukawa (2012) reconstructed museums with a trolley, assuming that all rooms are flat-floored and rectangular (even when they are not). Liu et al. (2010) modeled non-flat-floored indoor environments with a human-portable backpack, but the employed laser-image-fusion technique is yet limited to hallways only. Holenstein et al. (2011) reconstructed unstructured, large scale indoor environments (caves) with a 3D voxelized-volume-based approach that requires the model to be watertight and hence fails if an open sky or windows are present. Bosse et al. (2012) designed a spring-mounted 3-D range sensor that employs the laser data in a simultaneous localization and mapping (SLAM) -scheme. The latter two approaches are intriguing, since they employ only one 2D laser scanner that is rotated in a way so that the whole 3D environment is captured, in contrast to the previous two approaches that both employ three 2D scanners. Practically, the simpler and cheaper a solution is, the better it is. As a continuum to scanner method design, Elseberg et al. (2013) and Bosse and Zlot (2013) present probabilistic SLAM methods that are among other means applicable for scanner trajectory post-optimization, thus effectively improving the quality and precision of the entire acquired point cloud.

In this paper, we propose an approach to localization through theoretical mechanics. In particular, an experimental device is constructed to verify if localization succeeds for a (holonomic) 1D trajectory. Our device consists of a 2D laser scanner that is embedded between two round discs, effectively forming a rolling device. The scanner trajectory is reconstructed solely from the point cloud data. Two assumptions are required for this; that the device does not slip against the floor and that the floor is flat. No localization devices e.g. inertial or rotation measurement units are used. To our best

knowledge, every approach so far has employed at least an IMU. The proposed approach paves the road for a MLS paradigm, where a sole 2D laser scanner is employed.

2. Proposed approach

The localization of the laser data requires a successful reconstruction of the sensor's path of movement. This path is formally known as a time-dependent trajectory $j(t)$ with six degrees of freedom, namely, three from location and three from orientation. We write out

$$j(t) = [\,\theta(t), \psi(t), \phi(t), x(t), y(t), z(t)\,]^{\,T}, \tag{1}$$

where θ is the pitch, ψ is the roll, and ϕ is the yaw angle. Time is denoted by t. Without any reference coordinate system, the successful reconstruction of the trajectory requires that these degrees of freedom are eliminated. If this is done so that the coordinates become subject to the constraint

$$f(\theta, \psi, \phi, x, y, z, t) = 0, \tag{2}$$

where f is a bijective function, then the system is *holonomic* (see e.g. Bloch (2003)). Otherwise, the system is *non-holonomic*. The difference between these is that in the first case, the localization can be done knowing only the initial and the current state of the system. In the latter case, the trajectory reconstruction by path integration requires accurate measuring of the position all along the path.

Consider MLS platforms in general. In order to capture a 3D environment with a 2D laser scanner, the scanner must be rotated about at least one axis[1]. If the 2D scanner is mechanically attached on a wheel, at a radius R, so that it can only rotate about one axis of rotation, rotational degrees of freedom are reduced by two, i.e. ϕ and ψ are constant. Furthermore, if this wheel is mechanically connected onto another wheel on which the mobile platform rolls, the sensor's movement follows a trajectory

$$j(t) = F(j_p(t)\,), \tag{3}$$

where $j_p(t)$ is the platform trajectory and F is a bijective function.

[1]For example, Faro Focus 3D scanner in 2D helical mode, which was used in our experiments has a field of view of 300 degrees.

It is assumed that the platform trajectory j_p is continuous, and bounded to move on a 2D plane, i.e. on the floor. The platform trajectory is then a function of two time-dependent variables

$$j_p = j_p(\, r(t),\, \sigma(t)\,), \tag{4}$$

where $r(t)$ is the distance traveled and $\sigma(t)$ is the steering angle, i.e. the tangent of the trajectory at time t. Following the scope of this paper, it is assumed here that the platform moves dead straight, $\sigma(t) = const.$, $x = const.$, and that the sensor is mechanically attached to the rolling wheel, which does not slip against the floor. The solution for the sensor trajectory $j(t) = (x, y, z)(t)$ follows from the one for a contracted cycloid

$$\begin{cases} x = const. \\ y = R_0\theta + (R_0 - R_1)\sin\theta \\ z = R_0 + (R_0 - R_1)\cos\theta \\ \theta = \theta(t) \end{cases}, \tag{5}$$

where θ is the angle of scanner zenith in radians, R_0 is the radius of the cycloid, and R_1 is the scanner position on the radius, see Fig. 1 a). At the beginning, the zenith is pointing upwards, $\theta(t = 0) = 0$. Hence, $\theta(t)$ is the path parameter that describes the scanner trajectory, and obtaining it solves localization.

In order to to obtain $\theta(t)$ from the data, a following concept is proposed. Each time the 2D scanner is perpendicular towards the floor (PTF), $\theta(t) = \pi + 2\pi n$, $n = 0, 1, 2, ...$, the scanning distance reduces to minimum R_1. We call this a PTF-observation, and keep track of these occurrences in the laser data series obtaining a time series $\{t_1, t_2, t_3, ...\}$. We will discuss in the Results and Discussion section, how the PTF observation is used in determining the phase of the rolling sensor. Here, we note that the PTF observation is robust to error, since data points from a large field of view can be used to interpolate the floor point precisely below the sensor. Also, stochastic errors in PTF observations do not cumulate with time as long as the no-slip condition with the floor applies.

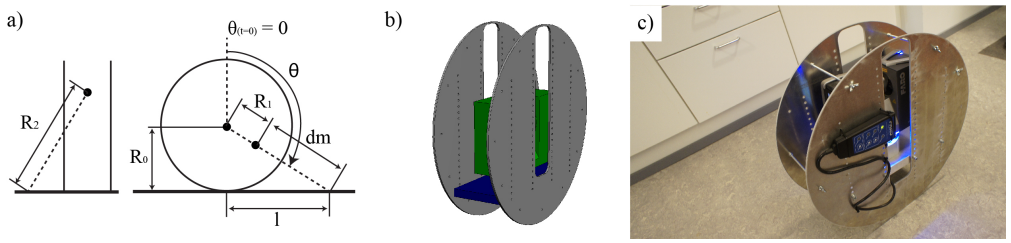

Figure 1: a) A schematic describing the path parameter $\theta(t)$ and constants R_0, R_1 and R_2. Measured distances are marked with d_m. b) VILMA assembly schematics. c) A 2D laser scanner, here Faro Focus 3D in helical mode, is attached in between two metallic discs so that it retains a wide viewing angle through the oval holes carved on the discs.

3. Concept Realization

3.1. The Build

The build of our experimental device, named VILMA, is depicted in Fig. 1. The radius of metal discs is $R_0 = 250$ mm. Faro Focus 3D laser scanner is mounted between the discs in helical 2D mode. 2D scans were conducted with 95 Hz frequency capturing 8534 points per a 300° field of view. A $1,5$ mm thick silicon ring padding was used to protect the scanner from tremor caused by dirt. With the padding, the distance covered with one rotation is 1572 mm. Based on this measure, the relative error in disc radius due padding elasticity is $\pm 0.1\%$.

Experiments were conducted both in a controlled laboratory environment i.e. a hallway, see Fig. 3 a), and in a rough environment, in an underground car park, see Fig. 3 b). The floor of the car park was covered with small particles, sand and dirt, and was sloped to direct water to the drains, implying that VILMA's altitude and rolling velocity change somewhat arbitrarily. VILMA was set to roll by a gentle manual push. When it started to slow down, another push was given, taking care that the rolling direction was not altered. We will discuss VILMA's dynamics later in detail in the Results section.

3.2. Path parameter

Based on experimentation, it was discovered that the concept of PTF-observations could be expanded so that the sensor angle θ is tracked continuously, except when the sensor turns parallel to the floor. For simplicity, VILMA was let to roll in one direction only, so that $\theta \in [0, \infty)$, and θ can be

assumed to grow monotonically. When the scanner records a distance onto a ground point, a triangle is formed between the recorded point, VILMA's central axis, and a point on the ground directly beneath the central axis, see Fig. 1 a). The scanner sits on the hypotenuse at a distance of $R_0 - R_i$ from VILMA's central axis. This triangle has one right angle, assuming that the floor is flat. Hence,

$$\theta = \arccos(R_0/d), \tag{6}$$

where $d = d_m + (R_0 - R_i)$, and d_m is the minimum measured distance over one full 2D circle observation.

Considering the minimum distance to the floor, the scanner's position on disk radius R varies between two values, R_1 and R_2, depending on whether the scanner is upside up or upside down. In the latter, i.e. PTF-position, the measured distance to floor is $R_1 = 13$ cm. In contrast, when the zenith of the sensor points up and the dead angle is under it, the measured distance to floor is $R_2 = 42$ cm. In order to know which value must be used for R, i.e. how to track the *phase change*, minimum distance is measured from specific scan angles. First, there is an about $22,5^\circ$ wide angle spanning equally about the scanner zenith, the *up channel*, and second, there are two about $12,25^\circ$ wide angles on each side of the dead angle, the *down channel*.

Two plots for d_m, determined separately from the up channel and the down channel data, are shown in Fig. 2. A phase change occurs when both channels record an augmenting value of d_m, say $d_m > 1.0$m. In other words, when the scanner is parallel to the floor. The fact, whichever channel records a smaller d_m, is used to determine VILMA's rolling phase, upside (zenith) up or upside down

$$\theta = \begin{cases} \arccos(R/d) & , d_{m,up} > d_{m,down} \\ \pi - \arccos(R/d) & , d_{m,up} < d_{m,down} \end{cases}. \tag{7}$$

Since θ increases monotonically, keeping track of phase changes as they occur in $\theta = \pi/2 + 2\pi n$ and $\theta = 3/2\pi + 2\pi n$ is rather straightforward. The roll angle is cumulated so that after each full round, $n \to n+1$, θ is increased by 2π. The cumulative roll angle $\theta(t)$, after noise-treatment that is discussed next, is shown in Fig. 2.

Noise in $\theta(t)$ is reduced by two steps. First, by enforcing the monotonic increase. Second, by identifying the parts where the angular acceleration $\ddot{\theta}(t)$

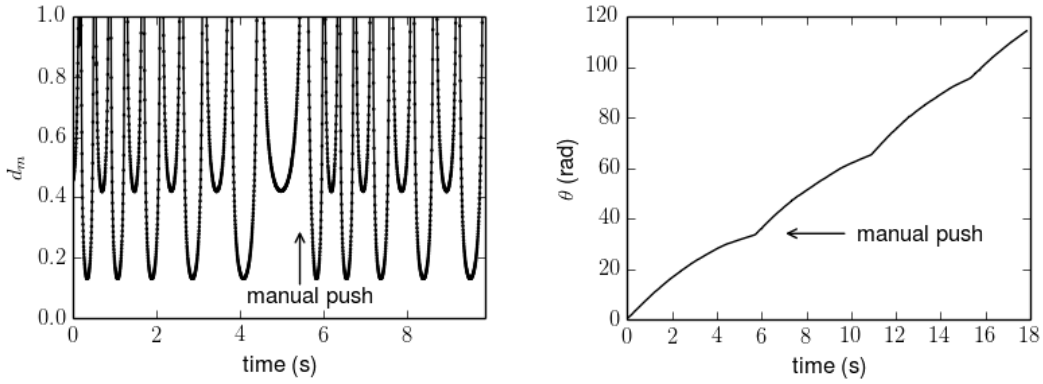

Figure 2: Left: d_m from up and down channels, with minimums of 0.13 m and 0.42 m, respectively. Only part of the data is shown for visual reasons. Right: Value of $\theta(t)$ determined from experimental data after noise treatment (see text). Note the different time scale in figures. Time is displayed in seconds. The approximate moment of one manual push is shown in both plots.

exceeds a threshold value of $\pi/16$, and then smoothing them out. For each such occurence, the remaining time-series of $\theta(t)$ is searched for stable sets that contain over 10 concurrent points with respective $\ddot{\theta}$'s under threshold value. The stable set with the lowest average value of $\theta(t)$ is selected, and the time-series is linearly interpolated to it, over the discrepancy.

Once the smoothed $\theta(t)$ is obtained, see Fig. 2, coordinate transformation for the sensor data (X, Z) is straightforwardly obtained from Eq. (5),

$$\begin{cases} x = X \\ y = R_0\theta + (R_0 - R_1)\sin\theta + \sin(\theta)\, Z \\ z = R_0 + (R_0 - R_1)\cos\theta + \cos(\theta)\, Z \end{cases} , \tag{8}$$

where (x, y, z) are the coordinates of the resulting 3D point cloud.

3.3. Point cloud filtering

Our point clouds have been filtered in two steps. First, to contain about 40% of its original points, so that most filtered points reside near PTF position. In particular, for each point $\sin(\theta)$ is added into a cumulative score S. If $S \geq 1$, the point is accepted and S is set to zero, otherwise the point is rejected. Second, the point cloud is filtered for noise with statistical outlier removal. Explicitly, for each point p the distance to 50 nearest neighbors is

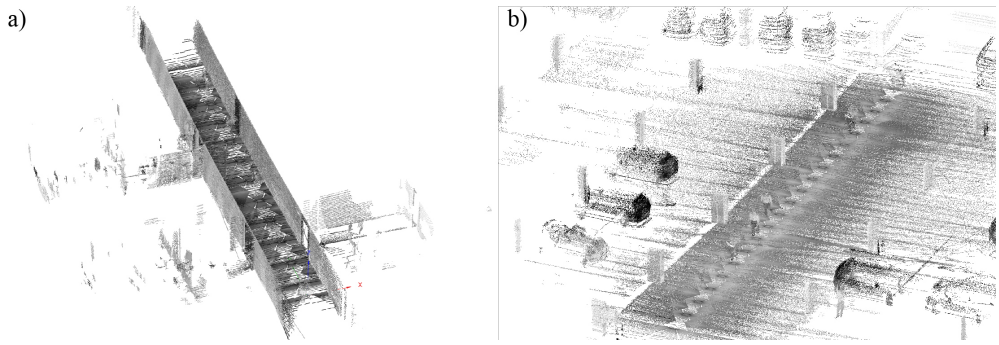

Figure 3: Results from VILMA. a) A simple hallway served as a benchmark for laboratory conditions b) The car park is visualized with an overviewing snapshot of the 3D point cloud with over 9 million points. The roof has been manually removed for visualization purposes. The greyscale colors in a) and b) depict the measured laser intensity.

computed to obtain a distribution, where μ_p and σ_p are the mean and the resulting standard deviation. Those points are rejected as outliers for which $\mu_p + \sigma_p > \mu + \sigma \cdot m_{sd}$, where μ and σ are the average mean and the standard deviation over all points, and $m_{sd} = 0.15$ is the standard deviation multiplier threshold selected manually for best visual outcome. Note that the point cloud generation is automatic, meaning that no manual work phases are required. The only exception to this was that we transferred the scanning data manually from the scanner to the computer, but this step is automatable e.g. through Faro Focus helical interface.

4. Results and Discussion

The 3D point cloud of a simple university hallway is shown in Fig. 3 a). The hallway was used in the concept realization phase as a benchmark for laboratory conditions.

Non-laboratory conditions were tested in a sloped-floor environment, see Fig. 3 b), and Fig. 4. Point cloud from VILMA that was obtained in some dozens of seconds was compared to the reference point cloud that was obtained in two hours. The latter was obtained from a combination of 8 conventional TLS scans using tripod-mounted Faro Focus 3D with spherical scan targets. There are two main types of distortion that affect the point cloud from VILMA. First, the sloped floor is reconstructed as flat in the resulting point cloud, implying that the slope of the floor is transferred onto the shapes of vertical objects e.g. pillars and walls, making them oblique,

Figure 4: Visual result comparison. a) Camera image from the scene b) Snapshot of the point cloud as obtained from VILMA after filtering explained in Section 3.3. In sloped-floor environment the resulting point cloud accuracy is reduced. Scanning is fast, taking some dozens of seconds. c) The reference point cloud taken with eight scans using a tripod and scan targets. Setting up scan targets, scanning and co-registering the point clouds took about two hours. The cars in a), b), and c) are at different locations because measurements were done at separate days.

see Fig. 4. The change in floor elevation along the trajectory is about 5 cm at length scales from 3 m to 9 m, see Fig. 5 b).

Second, the theoretically uniform rolling direction is impossible to keep unaltered with manual pushes. This results in the rotation of the point cloud coordinates and shows as displaced objects in the xy-plane, see Fig. 5 a). Experimentally, we found out that the transition perpendicular to the expected trajectory Δx can be kept with-in 0.1 meters for each traversed 10 meters, i.e. with-in reasonable bounds of $1:100$, in suitable conditions. In the chosen pathological sloped-floor environment deviation reaches $\Delta x \approx 0.5$ meters in 30 meters, see Fig. 5 b), due to conic-like shape of the floor curvature.

VILMA is designed for a flat floor. In order to develop a height correction model for sloped floors, the two variables in Eq. 6 would need to be decoupled. Installing an orientation sensor, which is beyond the scope of this paper, might be an appropriate solution, but a sufficient decoupling might also be achieved if the VILMA's dynamics were includable in the holonomic constraints of Eq. 5. To determine whether the latter is possible, even partially, the VILMA's angular velocity $\dot{\theta}$ is plotted in Fig. 6. The range of data is selected so that the moment of the same manual push shown in Fig. 2 is visible. Two representations of $\dot{\theta}$ are shown, the value shown with thin blue line is calculated with forward finite differences using three elements from time-series of θ, and the solid black line is a local average of this. Technically, an algorithm that takes the local average over 21 elements is ran twice. The solid black line displays normal deacceleration at long time scales, which is expected, but also shows that VILMA's movement is far

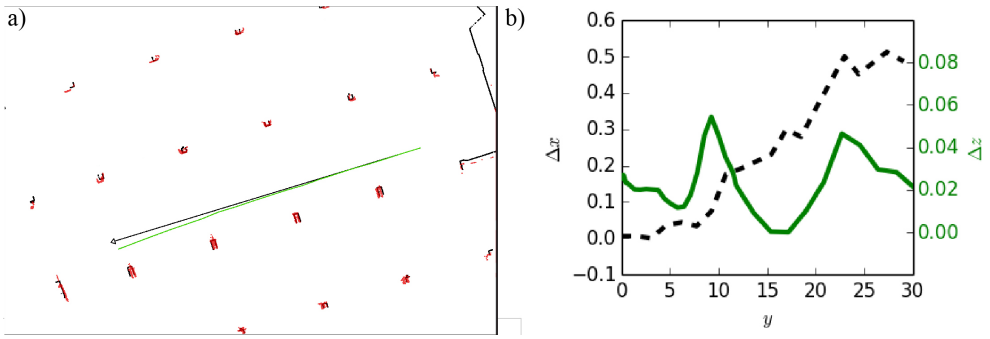

Figure 5: a) Match between the reference point cloud (grey) and the one measured with VILMA (red). Discrepancy between the results is due to platform rolling direction alteration. Traversed platform path of about 30 meters is shown with a solid green line. The path deviates from the straight black line drawn for visualization purposes, the deviation is elaborated in b). b) The deviation perpendicular to the assumed straight platform trajectory as a function of the rolled distance y, in meters. The deviation in x-direction (Δx) is shown with a dashed black line, and uses the left-side scale. The elevational deviation in z-direction (Δz) uses the right-side scale, and is shown with a green solid line. It demonstrates the physical irregularity of the floor sloping.

from steady. VILMA's dynamics is influenced by an inhomogeneous mass distribution around the rolling axis, which causes VILMA to accelerate and deaccelerate during rotation. Due to the arbitrariness of VILMA's acceleration, it is rather obvious that the holonomic constraints cannot be extended with $\dot{\theta}$. Despite this, it is remarkable how the proposed approach, and the results in Figs. 3, 4, and 5, are robust to VILMA's dynamics.

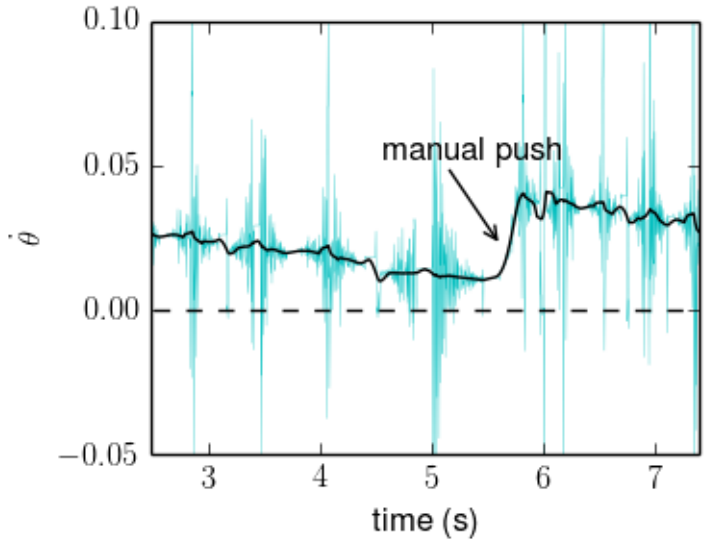

Figure 6: Angular velocity $\dot{\theta}$ as a function of time (in seconds), shown for a part of data calculated with finite differences (thin vertical blue lines), and then smoothed by local averaging (solid black line). Small variations follow e.g. from the unhomogeneous distribution of mass around VILMA's central axis. The steep upward slope is due to a manual push.

For final validation purposes, we consider the distances between the concrete pillars closest to the scanner path, see Fig. 5 a). The ground truth distance between consequent pillars is 781 ± 0.5 cm, measured at the bottom, from locations of faces whose normals point to the same direction. The respective measure is 779 ± 2 cm, obtained from a point cloud that is produced during the VILMA's first four rolling cycles, by taking a 25 cm high slice from just above the floor, and computing average locations of the pillar faces. Hence, the accuracy is $1 : 300$ of object size. It is then highly probable that the oblique surfaces in the point cloud can be corrected with automatic post-optimization trajectory-based methods such as in Elseberg et al. (2013) and Bosse and Zlot (2013). This is part of the future work, and outside the scope of this paper.

5. Conclusions

3D reconstruction typically requires that the laser scanning data is localized in a post-processing step. The basic paradigm is to obtain the laser

data from the scanner and the localization data, wholly or partly, from some other source, for example a combination of GNSS and IMU. When accurate localization data is not available, the reconstruction fails. In this paper, we propose an approach how this localization data can be replaced with information obtained solely from 2D laser scanning data. We call this *intrinsic localization*. In order to succeed, we use prior knowledge about the scanner device, called VILMA, and the environment, in two steps

1. The scanner trajectory is subject to holonomic constraints, assuming that VILMA rolls on a flat floor without slipping, and is therefore completely described by the path parameter, i.e. the scanner's rotation angle, θ.
2. The rotation angle θ can be extracted from the laser data by extending the proposed PTF-concept. PTF stands for the scanner being Perpendicular To the Floor. Finally, the scanner trajectory is solved by interpolating θ into a monotonically increasing and continuous form.

As results, 3D reconstructions from a flat-floored hallway and a sloped-floor car park are presented. VILMA's current design restricts the functionality on flat floors only, but future work aims to develop an appropriate height correction model to overcome this limitation. Furthermore, work to construct an advanced prototype to handle motion on a free 2D surface is also under way. For these, post-optimizing scanner trajectory by using place recognition in laser data is a promising direction.

Acknowledgements

The authors wish to thank Academy of Finland for funding this research, Grant No. 257755 (VL), and CoE-LaSR Grant No. 272195. Also, the support from the Finnish Funding Agency for Innovation, EUE project No. 2141226 is acknowledged.

References

Barber, D., Mills, J., Smith-Voysey, S., 2008. Geometric validation of a ground-based mobile laser scanning system. ISPRS Journal of Photogrammetry and Remote Sensing 63 (1), 128–141.

Bloch, A. M., 2003. Nonholonomic mechanics and control. Vol. 24. Springer.

Bosse, M., Zlot, R., 2009. Continuous 3d scan-matching with a spinning 2d laser. In: Robotics and Automation, 2009. ICRA '09. IEEE International Conference on. pp. 4312–4319, iD: 1.

Bosse, M., Zlot, R., 2013. Place recognition using keypoint voting in large 3d lidar datasets. In: Robotics and Automation (ICRA), 2013 IEEE International Conference on. pp. 2677–2684, iD: 1.

Bosse, M., Zlot, R., Flick, P., 2012. Zebedee: Design of a spring-mounted 3-d range sensor with application to mobile mapping. Robotics, IEEE Transactions on 28 (5), 1104–1119, iD: 1.

Elseberg, J., Borrmann, D., Nüchter, A., 2013. Algorithmic solutions for computing precise maximum likelihood 3d point clouds from mobile laser scanning platforms. Remote Sensing 5 (11), 5871–5906.
URL http://www.mdpi.com/2072-4292/5/11/5871

Holenstein, C., Zlot, R., Bosse, M., 2011. Watertight surface reconstruction of caves from 3d laser data. In: Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on. pp. 3830–3837, iD: 1.

Kukko, A., 2013. Mobile laser scanning - system development, performance and applications. Ph.D. thesis, Aalto University.
URL https://aaltodoc.aalto.fi/handle/123456789/12100

Kukko, A., Kaartinen, H., Hyyppä, J., Chen, Y., 2012. Multiplatform mobile laser scanning: Usability and performance. Sensors 12 (9), 11712–11733, 23112679.
URL http://www.mdpi.com/1424-8220/12/9/11712

Liu, T., Carlberg, M., Chen, G., Chen, J., Kua, J., Zakhor, A., 2010. Indoor localization and visualization using a human-operated backpack system. In: Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on. IEEE, pp. 1–10.

Thomson, C., Apostolopoulos, G., Backes, D., Boehm, J., 2013. Mobile laser scanning for indoor modelling. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-5/W2.

Xiao, J., Furukawa, Y., 2012. Reconstructing the world's museums. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (Eds.),

Computer Vision - ECCV 2012. Vol. 7572 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 668–681.
URL http://dx.doi.org/10.1007/978-3-642-33718-5_48