STRATEGIES TO INTEGRATE IMU AND LIDAR SLAM FOR INDOOR MAPPING

S. Karam [1, *], V. Lehtola [1], G. Vosselman [1]

[1] Dept. of Earth Observation Science, Faculty ITC, University of Twente, 7514 AE Enschede, The Netherlands
(s.karam, v.v.lehtola, george.vosselman)@utwente.nl

Commission I, WG I/7

KEY WORDS: Indoor Mapping, SLAM, IMU, Mobile Laser Scanning, LiDAR, 6DOF Pose Estimation, Point Clouds

ABSTRACT:

In recent years, the importance of indoor mapping increased in a wide range of applications, such as facility management and mapping hazardous sites. The essential technique behind indoor mapping is simultaneous localization and mapping (SLAM) because SLAM offers suitable positioning estimates in environments where satellite positioning is not available. State-of-the-art indoor mobile mapping systems employ Visual-based SLAM or LiDAR-based SLAM. However, Visual-based SLAM is sensitive to textureless environments and, similarly, LiDAR-based SLAM is sensitive to a number of pose configurations where the geometry of laser observations is not strong enough to reliably estimate the six-degree-of-freedom (6DOF) pose of the system. In this paper, we present different strategies that utilize the benefits of the inertial measurement unit (IMU) in the pose estimation and support LiDAR-based SLAM in overcoming these problems. The proposed strategies have been implemented and tested using different datasets and our experimental results demonstrate that the proposed methods do indeed overcome these problems. We conclude that IMU observations increase the robustness of SLAM, which is expected, but also that the best reconstruction accuracy is obtained not with a blind use of all observations but by filtering the measurements with a proposed reliability measure. To this end, our results show promising improvements in reconstruction accuracy.

1. INTRODUCTION & MOTIVATION

Indoor mapping is important for a wide range of applications, such as virtual tourism, facility management, interior design. Up-to-date 3D indoor maps of public buildings (hospitals, shopping malls, stations, airports, etc.) are also a prerequisite for navigation within these locales. Rapid advancements in light detection and ranging (LiDAR) technology, IMUs, optical instruments (cameras) have thus led to the development of many indoor mobile mapping systems (IMMSs).

The state-of-the-art IMMS consists of a movable platform equipped with laser scanners, IMUs and/or cameras to capture the indoor environment. Based on the selected moving platform, the developed IMMSs can be grouped into pushcart-based systems, such as Viametris i-MMS and iMS3D trolleys (Viametris, 2020), hand-held systems such as ZEB1 and ZEB REVO (Bosse et al., 2012; GeoSLAM, 2020) and backpack-based systems such as BIMAGE Backpack (Blaser et al., 2019) and Jafri et al. (2019). All these systems would solve SLAM algorithms for positioning indoors.

Unlike human-carried systems, pushcart-based systems do not have the ability to access whole interior areas, such as staircases. Thomson et al. (2013) have appraised the performance of ZEB1 and Viametris i-MMS by implementing two comparisons against a reference scan from the Faro Focus3D laser scanner. Through this, he found that ZEB1 is less compatible with the FARO cloud than i-MMS. Besides hardware, the software might also restrict the use of the IMMS. For example, Visual-based SLAM will fail in a textureless environment as it is based on matching similar features in consecutive images.

Karam et al. (2019a) have presented a backpack mobile mapping system that solves planar feature-based SLAM to obtain 3D point clouds of indoor environments. The 6DOF pose estimates are constrained by spline functions that guarantee a level of smoothness for the trajectory. The system consists of three laser range finders (LRFs - Hokuyo UTM-30LX scanners with 30 m range) that contribute to the 6DOF pose estimation of the system, and Xsens MEMS IMU is used for pose prediction (Karam et al., 2019b). As shown in Figure 1, the system is designed to have one scanner being horizontally positioned (S_0), while the other two scanners are slanted and positioned to the right and left of the horizontal one. The IMU is then mounted under the horizontal scanner. The LiDAR-based SLAM is designed to map indoor environments with planar and vertical structures through the linear segments detected in the single scanlines (Vosselman, 2014). Each laser point in a linear segment which is associated to a plane in the SLAM map forms an observation equation for the 6DOF pose estimation.
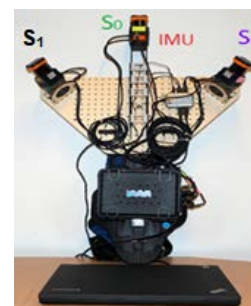

Figure 1. The used laptop and the current backpack system with four sensors mounted: three scanners S_0, S_1, and S_2 and Xsens MEMS IMU (below S_0)

This backpack system does not have multi-line LiDARs and therefore is in need of IMU support. This is because it is sensitive to a greater number of pose configurations where the geometry of laser observations is not strong or sufficient enough to reliably estimate the 6DOF pose of the system. Regardless of this sensitivity, the problem itself is generic. Even with multi-line LiDARs, pathological pose configurations can easily be found

---

* Corresponding author

(e.g long homogeneous hallways). Hence, this problem is present with all systems, motivating our research.

Weighting the balance between the inertial (or motion) measurements and the optical measurements is one of the key questions in SLAM post-optimization (Thrun and Montemerlo, 2006). Although it is well-known that weighing should be done, the exact way on how to do it is situational and often depends on the design of the sensor system and the environment. To this end, we shall compare two sensor fusion strategies: one that fuses all the data without considering the reliability of that data, and one that employs a proposed reliability measure to detect individual erroneous poses.

In this paper, we set out to study the above-mentioned strategies for IMU-SLAM integration. From Karam et al (2019b), it is obvious that the IMU can be used to guide the SLAM algorithm so that the algorithm avoids some pitfalls of the problematic measurement geometry. In contrast to the ZEB1 (Bosse et al., 2012) where the IMU is essential for the SLAM workability, the IMU plays a supportive role in our SLAM. However, the big question is whether the IMU observations can be used also in the (final) pose estimation so that (1) pathological pose configurations are overcome and (2) the effect of IMU's own drift to the pose estimates remains negligible. We answer to these very questions by proposing a combination of (1) a reliability measure for pose estimates derived from LiDAR observations and (2) coordinate and known-velocity updates, which reset the IMU drift. This resetting technique does not require a specific data capture mode to eliminate the IMU drifts as it is the case with the well-known Zero Velocity Updates (ZUPTs) technique (Chow et al., 2014).

The remainder of this paper is organized as follows: Section 2 presents the related works, while Section 3 describes our proposed strategies to integrate the IMU with LiDAR into SLAM after a short explanation of the IMU-less SLAM. In Section 4, we present results and discuss limitations of our approaches. The paper ends with conclusions in Section 5.

2. LITERATURE REVIEW

Studies show IMU integration increases the SLAM robustness regardless of the type of the optical sensor. Many works tend to use LiDAR-based SLAM algorithms that incorporate one or more scanners in the pose estimation. NavVis, for example, provided several solutions to map indoors as a trolley-based mobile mapping system that consisted mainly of scanners, IMUs and cameras (NavVis, 2019). GeoSLAM company evolved several versions of the hand-held ZEB-scanner (GeoSLAM, 2020) such as ZEB-REVO, ZEB-REVO RT, and ZEB Horizon. Besides the trolley-based and hand-held systems, there are several backpack systems (Lehtola et al., 2017). The Würzburg Backpack incorporates a 3D scanner (RIEGL VZ400), a 2D scanner (SICK LMS100) and an IMU that was utilized in the initial trajectory estimation (Lauterbach et al., 2015). Their experimental results showed maximum error about $7°$ in orientation and 25 cm in positioning. Zhang et al. (2017) attached an IMU to a LiDAR system to estimate odometry in real-time. All their walking experiments show that this combination improves the accuracy of motion estimation (See Table 3 in Zhang et al, 2017). The MEMS IMU in BIMAGE system, a component that combines two Velodyne VLP-16 laser scanners and one panorama camera, was exploited to estimate the camera orientation (Blaser et al., 2019). They solved 3D LiDAR-based SLAM for the Cartographer that combines the laser and IMU data, and then applied an image-based georeferencing approach

to improve the camera pose estimation. Chow et al. (2014) integrated a MEMS IMU in their stop-and-go mobile mapping system (Scannect) to support the vision-based localization in case the scene lacks features to be matched. They captured the data in a stop-and-go movement mode in order to utilize ZUPTs to eliminate their IMU drifts.

3. IMU-SLAM INTEGRATION STRATEGIES

As the backpack system is a mobile and multi-sensor system, we defined two main coordinate systems. The frame system (f) which is constantly moving. The data of all sensors are registered in this system. The model system (m) which is a fixed system and used to register the moving frame system over time. The final point clouds are defined in this system as well.

In our planar feature-based SLAM, we modelled the frame 6DOF pose parameters $(X, Y, Z, \omega, \phi, \kappa)$ in the model coordinate system (m) as functions of time using splines, and the planes were modelled by the normal vector and distance to the origin. The coefficients of the pose splines and the plane parameters are estimated simultaneously (Karam et al., 2019a). Thus, the adjustment process within SLAM does not only estimate and update the pose parameters but also the planes, which goes through different stages, as listed below. Figure 2 illustrates these stages exemplary. The differences between these stages are the splines' order, parameters to estimate, and the number of scans involved.

• **Local adjustment:** runs over a few successive scans captured during 0.1 s or slightly longer, and relies on the pose predicted by the IMU to check data association between the newly captured points and the previously reconstructed planes (Karam et al., 2019b). Each assigned point forms a laser observation equation and participates in the estimation of pose parameters that are modelled using linear splines. The laser observation equation is formulated based on the expectation that the distance between a point and its associated plane, i.e. plane that the point belongs to, equals zero. See Vosselman (2014) for details. The pose parameters are estimated and updated in this adjustment, while parameters of earlier instantiated planes are kept fixed.

• **IMU-based prediction:** runs at the beginning of each local adjustment to test the data association as mentioned above. Here, we utilized the strength of the IMU in short-term pose prediction and we predicted the pose within a time window that covers the time interval of one scan (25 ms). The IMU drifts are reset at the start time of each prediction window by using the position and the approximate velocity of the system estimated from the previous local adjustment at that time. Then, using the IMU data taken within this window, we predict the pose of the next scan (Karam et al., 2019b).

• **Section adjustment:** is executed when a plane has been observed for 0.5 s and runs over all successive scans that are captured in this period. The purpose of this adjustment is to improve the accuracy of the parameters of the plane instantiated 0.5 s ago. The increase in the number of laser observations enables us to use cubic splines instead of linear ones. The outputs of the local adjustment, including the estimated poses and instantiated planes, are inputs of this section adjustment as approximates. Consequently, the system poses, and planes states are estimated and updated in this stage.

• **Global adjustment:** combines all the captured scans in a final integral adjustment that provides an optimal estimate of all instantiated planes, along with a complete trajectory of the

system. The pose splines resulting from this adjustment are used to reconstruct the final point cloud in the model coordinate system.
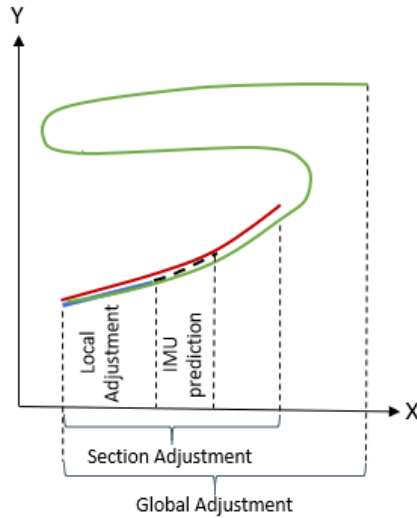

Figure 2. An exemplary representation of the prediction and adjustment processes within SLAM.

For the purpose of increasing the robustness of these adjustment processes against the aforementioned problematic areas, we developed the following three strategies for IMU-SLAM integration. They are comparable to Zhang et al. (2017) in terms of IMU integration purposes. In these strategies, the IMU participates not only in the pose prediction but also in the pose estimation.

3.1 IMU-SLAM Switching

The principle of the IMU-SLAM switching strategy is based on a replacement of the local adjustment-based pose by the IMU-based pose prediction in case the local adjustment is considered unreliable. The replacement occurs when the local adjustment fails or when the geometry of the laser observations is insufficient to estimate the pose. We decide whether or not to switch to the IMU-based estimate based on several indicators that measure the reliability of the SLAM-based estimate.

The first indicator is the reciprocal condition number (rcond) of the normal matrix which tells how close/far the normal matrix is to being singular. A very small rcond indicates that the normal matrix is close to being singular or badly scaled. Another indicator of a poor conditioned equation system is a high correlation between the estimated parameters. We determined the correlation matrix and check the maximum absolute value of all non-diagonal elements. If this value is approaching 1 for any pair of parameters, these parameters are highly correlated and it would, therefore, be hard to determine them separately.

In addition, we added a well-known scan matching technique, iterative closest points (ICP), to the proposed method. This technique is commonly used as a pose estimation method to support SLAM by defining the relative transformation between successive scans (Lee et al., 2011). We utilized the strength and efficiency of 2D ICP in estimating the relative 2D transformation (X, Y, κ) when it matches two successive scans of the horizontally mounted scanner. As the time interval of one scan is 25 ms, a large rotation between two scans in this short period is not expected. However, we would not rely on the 3DOF ICP-based pose to move forward in SLAM. Rather, we are simply determining the differences in the estimated pose parameters (X,

Y, k) between the ICP and local adjustment. A big difference would raise the suspicion that one of the two methods is wrong, and this is used as an indicator to switch to the IMU-based prediction.

3.2 IMU-based Pose Estimation

The difference with the switching technique is that we always use the IMU-based pose prediction instead of the local adjustment-based pose, no matter whether the local adjustment seems reliable or not. The role of the local adjustment is limited to instantiate planes and for the checking of data association.

3.3 IMU-SLAM Joint Estimation

In this strategy, we deepen the SLAM-IMU integration by including the IMU observations in the 6DOF pose estimation besides the laser observations. Our Xsens MEMS IMU measures three-dimensional angular velocity and three-dimensional dynamic acceleration over time. Some of our IMU specifications are listed in Table 1. Similar to Hussnain et al. (2018), we formulate six IMU observation equations at each timestamp, t, as described in the following subsections.

SLAM	Gyroscope	Accelerometer
Bias repeatability	$0.2°/s$	0.03 m/s^2
In-run bias stability	$10°/h$	40 µg

Table 1. Key specifications of the Xsens MEMS IMU

3.3.1 Acceleration Observation Equations

The IMU accelerometers observe the accelerations $\ddot{T}_{imu}^s = (\ddot{X}_{imu}^s, \ddot{Y}_{imu}^s, \ddot{Z}_{imu}^s)$ in the sensor coordinate system (s) which needs to be rotated to the frame coordinate system (f) with the time-independent rotation matrix $R_s^f = R_z(90)$. This will also need to be performed on the model system (m) with the time-dependent rotation matrix $R_f^m(\omega, \varphi, \kappa)$. Hence,

$$\ddot{T}_{imu}^m = R_f^m \, R_s^f \, \ddot{T}_{imu}^s \qquad (1)$$

As the resulting accelerations $(\ddot{T}_{imu}^m)$ should correspond to the second-order derivative of the backpack's location $(\ddot{T}_f^m)$ in the model system, the following observation equations are formulated:

$$\ddot{T}_f^m = R_f^m \, R_s^f \, \ddot{T}_{imu}^s - \begin{pmatrix} 0 \\ 0 \\ g \end{pmatrix} \qquad (2)$$

where $\ddot{T}_f^m = (\ddot{X}_f^m, \ddot{Y}_f^m, \ddot{Z}_f^m)$.

Since the z-axis in the model system is assumed to be vertical, we compensate for the effect of gravity in the accelerometer reading by subtracting the average gravitational acceleration (g) from the acceleration along this axis.

As the pose parameters are modelled using splines in the laser observation equations, we also modelled the accelerations using splines. Splines are polynomial functions and it is straightforward to derive the accelerations $(\ddot{T}_f^m)$ as the second-order derivatives of the translations (T_f^m). For example, for the translation X spline $X(t) = \sum_i \alpha_{x,i} B_i(t)$, the acceleration $\ddot{X}$ spline becomes $\ddot{X}(t) = \sum_i \alpha_{x,i} \ddot{B}_i(t)$, where $\alpha_{x,i}$ is the X spline coefficient to be

estimated on interval i. Hence, both the translations and accelerations are expressed in terms of the same to-be-determined spline coefficients.

By doing the same for the measured accelerations along other axes, we formulated the following linearized acceleration observation equations in which the upper index ° refers to the approximate values:

$$\ddot{T}_f^{m0} - R_f^{m0} R_s^f \ddot{T}_{imu}^s + \begin{pmatrix} 0 \\ 0 \\ g \end{pmatrix} \qquad (3)$$

$$= - \begin{pmatrix} \sum_i \Delta\alpha_{x,i} \ddot{B}_i \\ \sum_i \Delta\alpha_{y,i} \ddot{B}_i \\ \sum_i \Delta\alpha_{z,i} \ddot{B}_i \end{pmatrix}$$

$$+ \frac{\partial R_f^{m0}}{\partial \omega} R_s^f \ddot{T}_{imu}^s \sum_i \Delta\alpha_{\omega,i} B_i$$

$$+ \frac{\partial R_f^{m0}}{\partial \varphi} R_s^f \ddot{T}_{imu}^s \sum_i \Delta\alpha_{\varphi,i} B_i$$

$$+ \frac{\partial R_f^{m0}}{\partial k} R_s^f \ddot{T}_{imu}^s \sum_i \Delta\alpha_{\kappa,i} B_i$$

where $\Delta\alpha_{x,i}, \Delta\alpha_{y,i}, \Delta\alpha_{z,i}, \Delta\alpha_{\omega,i}, \Delta\alpha_{\varphi,i}$ and $\Delta\alpha_{\kappa,i}$ are the unknown increments of the pose splines coefficients.

In the first iteration of the estimation process, $\ddot{T}_f^{m0}$ and R_f^{m0} are derived from the SLAM as the approximate acceleration and rotation of the system, respectively. This will reset the IMU's accelerometers drift which helps to mitigate the effects of the IMU biases (see Table 1) on the pose estimation.

3.3.2 Angular Velocity Observation Equations

Similarly, the angular velocity observation equations were formulated as the IMU angular velocities $\dot{V}_{imu}^s = (\dot{\omega}_{imu}^s, \dot{\varphi}_{imu}^s, \dot{\kappa}_{imu}^s)$ should be related to the first-order derivatives of the backpack's rotation angles $\dot{V}_f^m = (\dot{\omega}_f^m, \dot{\varphi}_f^m, \dot{\kappa}_f^m)$. However, while the IMU gyroscopes observed the angular velocities around the axes of the IMU sensor frame, the backpack rotation angles are defined around the axes of the model system and are used to rotate from the backpack frame to the model system. Therefore, in order to determine the relationship between these two groups of angular velocities, we first need to define the direction and order of rotation. The rotation from the model system to the IMU sensor frame system can simply be defined as the inverse of the rotation $(R_f^m R_s^f)$:

$$R_m^s = (R_f^m R_s^f)^T = R_f^s R_m^f = R_f^s R_3(\kappa_f^m)^T R_2(\varphi_f^m)^T R_1(\omega_f^m)^T \quad (4)$$

As κ is the first rotation applied when rotating from the backpack frame to the model system, the angular velocity around the z-axis ($\dot{\kappa}_f^m$) would not have to be rotated by R_m^f (Karam et al., 2019b). Indeed the IMU angular velocity $\dot{\varphi}_{imu}^s$ around the y-axis does not hold a direct correspondence with the first derivative of φ_f^m when simply rotating (R_f^s) from the backpack frame to the IMU sensor system. This is as the y-axis has already been rotated by $-\kappa_f^m$ around the z-axis prior to the measuring of $\dot{\varphi}_{imu}^s$ in the

IMU sensor system. As such, the derivative of φ_f^m should also be rotated to the frame system. Similarly, the derivative of ω_f^m would need to be rotated by $-\varphi_f^m$ around the y-axis, and by $-\kappa_f^m$ around the z-axis, in order to obtain an angular velocity vector in the frame system. Hence, after rotating from the frame system to the IMU system, we were able to obtain the angular velocities as defined by the pose splines in the IMU system. This leads to the following observation equations:

$$\begin{pmatrix} \dot{\omega}_{imu}^s \\ \dot{\varphi}_{imu}^s \\ \dot{\kappa}_{imu}^s \end{pmatrix} = R_f^S \begin{pmatrix} 0 \\ 0 \\ \dot{\kappa}_f^m \end{pmatrix} + R_f^S R_3(\kappa_f^m)^T \begin{pmatrix} 0 \\ \dot{\varphi}_f^m \\ 0 \end{pmatrix}$$

$$+ R_f^S R_3(\kappa_f^m)^T R_2(\varphi_f^m)^T \begin{pmatrix} \dot{\omega}_f^m \\ 0 \\ 0 \end{pmatrix} \qquad (5)$$

It can be shortened to:

$$\dot{V}_{imu}^s = R_f^S S_m^f \dot{V}_f^m \qquad (6)$$

with $S_m^f = \begin{pmatrix} \cos\varphi\cos\kappa & \sin\kappa & 0 \\ -\cos\varphi\sin\kappa & \cos\kappa & 0 \\ \sin\varphi & 0 & 1 \end{pmatrix}_m^f$

where S_m^f is the transformation matrix from the model to the frame system.

We model the angular velocities using splines as well by taking first-order derivatives of the rotation angles. For instance, for the rotation angle ω spline $\omega(t) = \sum_i \alpha_{\omega,i} B_i(t)$, the angular velocity $\dot{\omega}$ spline becomes $\dot{\omega}(t) = \sum_i \alpha_{\omega,i} \dot{B}_i(t)$, where $\alpha_{\omega,i}$ is the ω spline coefficient to be estimated on interval i.

Hence, the linearized angular velocity observation equations becomes:

$$\dot{V}_{imu}^s - R_f^s S_m^{f\,0} \dot{V}_f^{m0} \qquad (7)$$

$$= R_f^s S_m^{f\,0} \begin{pmatrix} \sum_i \Delta\alpha_{\omega,i} \dot{B}_i(t) \\ \sum_i \Delta\alpha_{\varphi,i} \dot{B}_i(t) \\ \sum_i \Delta\alpha_{\kappa,i} \dot{B}_i(t) \end{pmatrix}$$

$$+ R_f^s \frac{\partial S_m^{f\,0}}{\partial k} \dot{V}_f^{m0} \sum_i \Delta\alpha_{k,i} B_i$$

$$+ R_f^s \frac{\partial S_m^{f\,0}}{\partial \varphi} \dot{V}_f^{m0} \sum_i \Delta\alpha_{\varphi,i} B_i$$

Similar to the acceleration equation, the angular velocities $\dot{V}_f^{m0}$ and the transformation matrix $S_m^{f\,0}$ in the first iteration of the estimation process are derived from the approximate splines. This helps to reset the IMU's gyroscopes drift and this, in turn, eliminates the attitude drift considerably.

3.3.3 Joint Estimation

For the joint estimation method, the IMU observation equations (eq. 3 and eq. 7) are added to the laser observation equations in all adjustment processes addressed above: local, section and global. This fusion enables us to use cubic splines instead of linear ones in the local adjustment. As our backpack system was mounted with three Hokuyo scanners with a scanning frequency of 40 Hz and 1080 points per scan line, every 25 ms the system

records 3240 laser points. The accelerations and angular velocities are recorded by the IMU with a sampling frequency of 200 Hz; thus, for a local adjustment within 0.1 s, we have 12960 laser points, 60 IMU acceleration readings, and 60 IMU angular velocity readings. However, not all laser points were assigned to planes and can contribute to the pose estimation, as this is based on the data association criteria.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In order to test the performance of the three proposed strategies for a SLAM-IMU integration, several experiments on different indoor environments were conducted. Three datasets were collected at the Institute of Geodesy and Photogrammetry building at the University of Braunschweig, Germany. They are denoted Diemen0, Diemen1, and Diemen2, respectively. The number of captured laser points and the walking duration for all test areas are approximately 38 million/5-minute, 73 million/9-minute, and 38 million/5-minute, respectively. Diemen0 is captured in the relatively cluttered basement, while the other two datasets are captured on the ground floor with different operators and trajectories. The mapped areas show distinct office environments that have many doors and windows. There were also several narrow rooms with glass windows and uneven curtains. In order to test for the aforementioned weaknesses, some doors were wide open, and others were opened by the operator while the data was being captured. As the rooms in Diemen1 and Diemen2 were nearly empty due to renovation works, the number of planar structures is limited. This, in turn, affected the estimability of the system pose using only laser observations. The previously mentioned features of the test areas sometimes constitute as obstacles, and this may cause some of previously developed LiDAR-based SLAMs to fail before the mapping was completed, see Table 3.

We mapped these test areas using our backpack indoor mobile mapping system and ran five different versions of SLAM algorithms on each dataset. For the IMU-SLAM Switching, the rcond and correlation thresholds are experimentally determined and set to 0.02 and 0.7, respectively. The differences thresholds with ICP to raise the suspicion in SLAM performance are set to 1 cm for (x, y), and 0.5° for κ.

For the purpose of simplicity, we use the following terminology for the five compared SLAM strategies;

LiDAR-based SLAM: relies on the linear extrapolation for pose prediction instead of the IMU (Karam et al., 2019a). **LiDAR-based SLAM with IMU prediction**: takes advantage of the IMU to predict the next pose (Karam et al., 2019b). **IMU-SLAM Switching, IMU-based Pose Estimation,** and **IMU-SLAM Joint Estimation** are the three IMU-SLAM integration techniques proposed in this paper (see Table 2).

SLAM Algorithm	6DOF Pose prediction	6DOF Pose estimation
LiDAR-based SLAM (Karam et al., 2019a)	LiDAR	LiDAR
LiDAR-based SLAM (Karam et al., 2019b)	LiDAR+IMU	LiDAR
Three methods proposed in this paper	LiDAR+IMU	LiDAR+IMU

Table 2. Comparison between the proposed methods in this paper and previously developed methods regarding the data source for pose prediction and estimation.

The most crucial aspect in the testing of the proposed methods was to check the robustness against poor laser observations geometry. Indeed, the improvement in terms of robustness was evident through the ability to handle more datasets. This is particularly evident for the Diemen1 data, which the LiDAR-based SLAM, even with the help of IMU prediction, failed to process, as shown in Table 3. While for the Diemen0 data, all methods that estimate the pose without the use of IMU observations failed. This failure is caused by the divergence of the global adjustment because of insufficient laser observations in a sequence of five or more intervals. Figure 3 shows an obvious case where the LiDAR-based SLAM fails without the support of an IMU. There is an insufficient amount of laser observations to estimate the translation along the Y-axis. Therefore, the system slides along the Y-axis and leads to an erroneous map. One reason for the poor laser observations is that the narrow wall in front of the operator has a large transparent object (glass window) which leads to missing or incorrect range measurements. Problem areas also include a non-flat panel radiator underneath the window and a winding curtain reflect in sparse laser scans in which it is hard to detect linear segments. There are also no observations on the wall at the opposite side of the room because of the 270° opening angle of the Hokuyo scanners.

In comparison to other methods, the IMU-SLAM joint estimation is the most robust method due to its ability to handle all datasets in this paper. The reliance on IMU allows the SLAM to reconnect to planes seen some longer time before; thus, this prevents the system from sliding in any direction and supports the SLAM in going ahead.
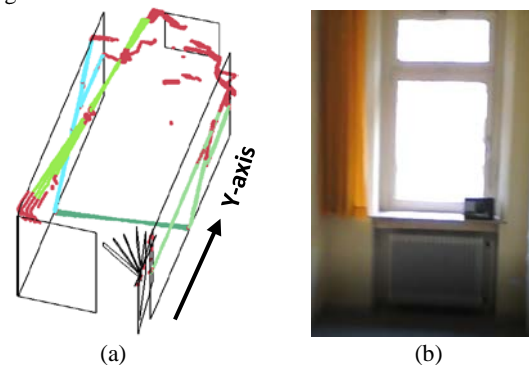

| (a) | (b) |

Figure 3. An example of a problematic area for the LiDAR-based SLAM. a) 3D reconstructed planes (black) with the assigned laser points from four scans (colours indicate point associations to a particular plane) and the points that are not associated to a plane (red). The arrow refers to the Y-axis direction in the model system. b) The wall is in front of the operator.

Figure 4 shows the generated point clouds of all test areas with colours showing the plane association. The points that are not associated to a plane and the points on the ceiling are removed for visualisation purposes.

We performed a comparison between the resulting point clouds based on several factors, such as the number of points assigned to the reconstructed planes and the root mean square error (RMSE) of the residuals, as demonstrated in Table 3. In this paper, the term residuals refers to the distances between the model planes and the points associated to them. Table 3 demonstrates that the use of IMU in pose estimation has increased significantly the number of assigned points to planes in Diemen2 dataset from about 24 million points to 28 million points. Approximately 97% of the assigned points have residuals below 3 cm which is 10% higher than the case of not using the

IMU. Also, the RMSE of the residuals decreased from 1.8 cm with LiDAR-based SLAM to about 1.3 cm. Consequently, the proposed strategies for IMU integration lead into a more robust data association between the captured points and the reconstructed planes.

As the SLAM-IMU joint estimation tries to fit the trajectory to the IMU observations, it is slightly sacrificing the minimization of the point-to-plane distances. Therefore, table 3 shows that this method is not always the best performing method regarding the RMSE of residuals, even if it is the most robust method.

Figure 5 compares the histograms of the points' residuals of all resulting point clouds for the Diemen2 dataset as it is the only dataset that was processed successfully by all methods. It demonstrates that approximately 57% of the assigned points have residuals below 1 cm when the IMU was not used at all. This improved to 61% when the switching technique took place, to 64% when the IMU was utilized in pose prediction and to about 70% when IMU contributes to the pose estimation. On the other hand, approximately 13% of the residuals exceeded 3 cm when the IMU was not used. This percentage is negligible when the IMU is utilized for pose prediction and estimation. As such, this is why the RMSE value of the assigned points' residuals decreases (Table 3).

However, the previous measurements, the number of assigned points and the RMSE of their residuals, do not adequately reflect the overall quality of the methods' performance. The lower RMSE indicates that we do a better fit of points to planes, but this does not necessarily mean that we have a better model. In addition, the RMSE could be influenced badly by an incorrect merge of two planes.

Therefore, we applied further quality measures which utilized the architectural constraints of walls (perpendicularity and parallelism) as they are predominant characteristics in the scanned areas (Karam et al., 2018). We assessed the ability of the system to capture the true geometry of the scanned environment. The results in Table 4 show the reconstruction accuracy with the IMU contribution is much better than without IMU. The percentages of small angles' errors ($<0.5°$) increase and of outliers ($>1°$) decrease in all proposed methods that utilize the IMU.
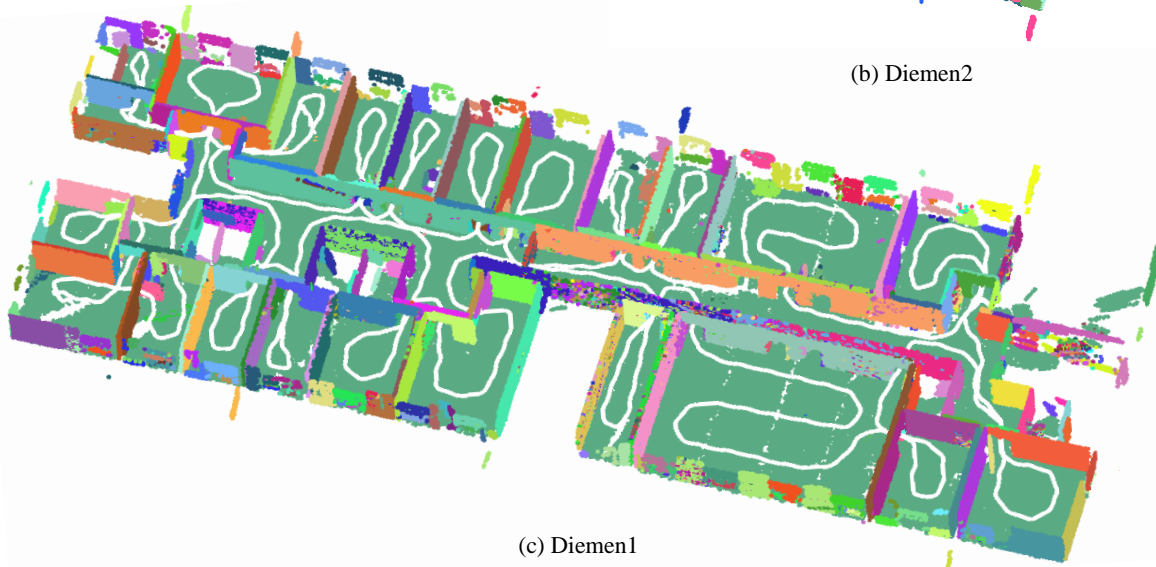
It also demonstrates that the reliance on the IMU at many locations did not negatively affect the internal consistency of walls; thus, it did not impede upon the correctness of the final 3D reconstruction. However, Table 4 shows that the best results were reached when the SLAM kept switching to the IMU.
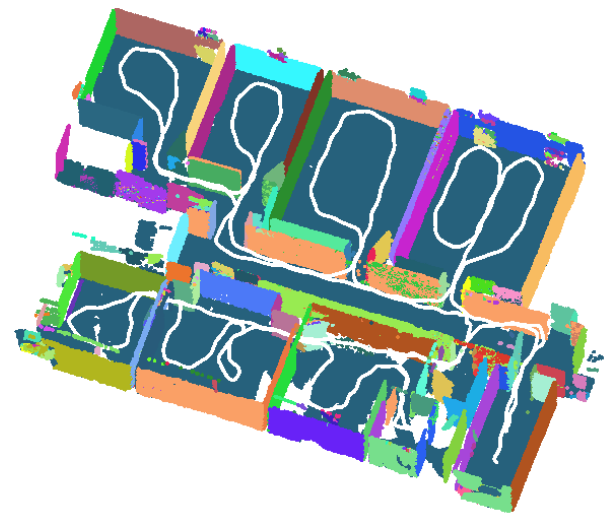
(a) Diemen0

(b) Diemen2

(c) Diemen1

Figure 4. The generated point cloud of the test areas with colours show plane association and trajectory followed (white). For visualization purposes, the points that are not associated to a plane and the points on the ceiling are removed from the point clouds.

Dataset / SLAM		Diemen0	Diemen1	Diemen2
LiDAR-based SLAM (Karam et al., 2019a)	Nr (million)			24
	RMSE (cm)	x	x	1.8
	Perc (%)			87
LiDAR-based SLAM (Karam et al., 2019b)	Nr (million)			24
	RMSE (cm)	x	x	**1.3**
	Perc (%)			96
IMU-SLAM switching	Nr (million)		**53**	**28**
	RMSE (cm)	x	**1.5**	1.4
	Perc (%)		96	97
IMU-based Pose Estimation	Nr (million)		53	**28**
	RMSE (cm)	x	1.8	1.4
	Perc (%)		96	97
IMU-SLAM Joint Estimation	Nr (million)	**28**	**53**	**28**
	RMSE (cm)	**1.5**	1.6	**1.3**
	Perc (%)	**98**	96	97

Table 3. Comparison between the performance of the proposed method on all datasets regarding the number of assigned points (Nr), the corresponding RMSEs of the residuals, and the percentage of residuals below 3 cm (Perc). The results of two previously developed methods are also listed at the beginning of the table for comparison. The letter "x" refers to the failure of the method to process the corresponding dataset.

Dataset / SLAM		Diemen0		Diemen1		Diemen2	
		perp	par	perp	par	Perp	par
LiDAR-based SLAM (Karam et al., 2019a)	[0° 0.5°[					58	50
	[0.5° 1°]	x		x		26	12
	[1° 5°]					16	38
	RMSE (deg)					1.34°	1.76°
LiDAR-based SLAM (Karam et al., 2019b)	[0° 0.5°[					**70**	66
	[0.5° 1°]	x		x		23	13
	[1° 5°]					7	21
	RMSE (deg)					0.66°	1.39°
IMU-SLAM Switching	[0° 0.5°[			69	62	67	65
	[0.5° 1°]	x		19	16	23	16
	[1° 5°]			12	22	10	19
	RMSE (deg)			**0.54°**	**1.31°**	0.96°	1.50°
IMU-based Pose Estimation	[0° 0.5°[			68	54	**70**	**71**
	[0.5° 1°]	x		17	16	23	11
	[1° 5°]			15	30	7	18
	RMSE (deg)			0.66°	1.36°	**0.63°**	**1.06°**
IMU-SLAM Joint Estimation	[0° 0.5°[	62	39	67	57	67	61
	[0.5° 1°]	25	30	18	16	23	12
	[1° 5°]	13	31	15	27	10	27
	RMSE (deg)	0.63°	1.34°	0.70°	1.45°	0.82°	1.46°

Table 4. Comparison between the performance of the proposed method on all datasets regarding the architectural constraints. The percentages of angles' errors in three different ranges for parallelism and perpendicularity and the corresponding RMSEs. The results of two previously developed methods are also listed at the beginning of the table for comparison. The letter "x" refers to the failure of the method to process the corresponding dataset.
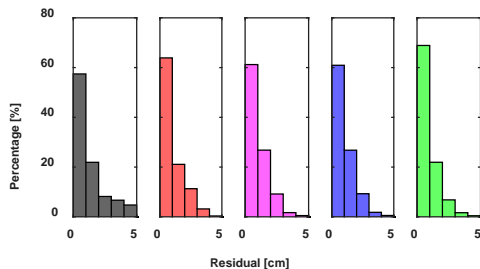

Figure 5. Histograms of the points' residuals of all resulting point clouds for Diemen2 dataset and for all methods, LiDAR-based SLAM (black), with IMU prediction (red), IMU-SLAM switching (magenta), IMU-based Pose Estimation (blue), and the joint estimation (green).

Overall, we observe an improvement in results when the erroneous poses that originate from LiDAR-based SLAM are replaced by the poses integrated from IMU measurements. This goes hand-in-hand with the general understanding that the less trustworthy measurements should receive no weight in SLAM optimization.

The proposed SLAM strategies are limited to (indoor) environ-

ments with planar structures. The IMU drift in velocity and heading is reset in each prediction or estimation loop (see Section 3), but the IMU biases are not auto-calibrated during scanning. This did not cause immediate problems in our experiments but should be taken into account with lower quality IMUs. In this case, this limitation could be addressed, e.g., for the gyro biases (Hyyti and Visala, 2015).

5. CONCLUSION

We have described several strategies to integrate the IMU with the LiDAR-based SLAM. Here, the IMU participates in the pose prediction and estimation. We conclude that the IMU integration improves the robustness of both the data association and pose estimation and therefore it is beneficial in the proposed SLAM approach. This improvement in the SLAM robustness expands the scope of application of our backpack mobile mapping system.

In future works, we intend to investigate whether the IMU is capable to generate a reliable prediction for a longer exposure period. Reliable prediction for a wider period will help to improve the hypothesis generation of planar structures which, in turn, enables the SLAM to sense non-vertical walls or non-horizontal floors/ceilings. Consequently, this would expand the applicability of our system.

ACKNOWLEDGEMENTS

REFERENCES

Blaser, S., Nebiker, S., Wisler, D., 2019. Portable image-based high performance mobile mapping system in underground environments - system configuration and performance evaluation. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 4, 255–262. doi.org/10.5194/isprs-annals-IV-2-W5-255-2019

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

Chow, J.C.K., Lichti, D.D., Hol, J.D., Bellusci, G., Luinge, H., 2014. IMU and multiple RGB-D camera fusion for assisting indoor stop-and-go 3D terrestrial laser scanning. *Robotics* 3, 247–280. doi.org/10.3390/robotics3030247

GeoSLAM Ltd. The ZEB-REVO Solution. Available online: https://geoslam.com/solutions/zeb-revo/ (accessed on 19.01.2020).

Hussnain, Z., Elberink, S.O., Vosselman, G., 2018. An automatic procedure for mobile laser scanning platform 6Dof trajectory adjustment. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 42, 203–209. doi.org/10.5194/isprs-archives-XLII-1-203-2018.

Hyyti, H., & Visala, A. (2015). A DCM based attitude estimation algorithm for low-cost MEMS IMUs. *International Journal of Navigation and Observation*, 2015. https://doi.org/10.1155/2015/503814

Jafri, S.R.U.N; Amjad, Hira; Sultan, Moazza; Khan, H.., 2019. Low cost 2D laser scanner based indoor mapping and classification system. *Int. Conf. Robot. Autom. Ind. (ICRAI), Rawalpindi, Pakistan.* 1–6. doi.org/doi: 10.1109/ICRAI47710.2019.8967399

Karam, S., Lehtola, V., Vosselman, G., 2019b. Integrating A low-cost MEMS IMU into a laser-based SLAM for indoor mobile mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences: 6th International Workshop LowCost 3D – Sensors, Algorithms, Applications. editor / P. Grussenmeyer ; A. Murtiyoso ; H. Macher ; R. Assi. Vol. XLII-2/W17 Strasbourg : International Society for Photogrammetry and Remote Sensing (ISPRS).* pp. 149-156. doi.org/10.5194/isprs-archives-XLII-2-W17-149-2019

Karam, S., Peter, M., Hosseinyalamdary, S., Vosselman, G., 2018. An evaluation pipeline for indoor laser scanning point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* IV-1, 85–92. doi.org/10.5194/isprs-annals-IV-1-85-2018

Karam, S., Vosselman, G., Peter, M., Hosseinyalamdary, S., Lehtola, V., 2019a. Design, calibration, and evaluation of a backpack indoor mobile mapping system. *Remote Sens.* 11. doi.org/10.3390/rs11080978

Lauterbach, H.A., Borrmann, D., Heß, R., Eck, D., Schilling, K., Nüchter, A., 2015. Evaluation of a backpack-mounted 3D mobile scanning system. *Remote Sens.* 7, 13753–13781. doi.org/10.3390/rs71013753

Lee, H-C., Lee, S-H., Lee, S-H., Lee, T-S., Kim, D-J., Park, K-S., Lee, K-W., Lee, B-H, 2011. Comparison and analysis of scan matching techniques for cooperative-SLAM. *Ubiquitous Robot. Ambient Intell. (URAI), 8th Int. Conf. on. IEEE.* 165–168.

Lehtola, V. V., Kaartinen, H., Nüchter, A., Kaijaluoto, R., Kukko, A., Litkey, P., Honkavaara, E., Rosnell, T., Vaaja, M.T., Virtanen, J.P., Kurkela, M., El Issaoui, A., Zhu, L., Jaakkola, A., Hyyppä, J., 2017. Comparison of the selected state-of-the-art 3D indoor scanning and point cloud generation methods. *Remote Sens.* 9, 1–26. https://doi.org/10.3390/rs9080796

NavVis. Available Online: https://www.navvis.com/ (accessed 20.12.2019).

Thomson, C., Apostolopoulos, G., Backes, D., Boehm, J., 2013. Mobile laser scanning for indoor modelling. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* vol. 2, part. 5/W2, 11-13 November, Antalya, Turkey 289–293.

Thrun, S., Montemerlo, M., 2006. The Graph SLAM algorithm with applications to large-scale mapping of urban structures. *Int. J. Rob. Res.* 25, 403–429. doi.org/10.1177/0278364906065387

Viametris iMS3D - iMS2D. Available Online: http://www.viametris.com/products/ims3d/ (accessed 15.01.2020).

Vosselman, G., 2014. Design of an indoor mapping system using three 2D laser scanners and 6 DOF SLAM. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 2.3 173.

Zhang, Ji, and Sanjiv Singh. "Low-drift and real-time lidar odometry and mapping." *Autonomous Robots* 41.2 (2017): 401-416.