A Laser-Aided Inertial Navigation System (L-INS) for Semi-structured Indoor Environments

Joel A. Hesch, Faraz M. Mirzaei, Gian Luca Mariottini, and Stergios I. Roumeliotis

Abstract— **This paper presents a novel 3D indoor Laser-aided Inertial Navigation System (L-INS) for the visually impaired. An Extended Kalman Filter (EKF) fuses information from an Inertial Measurement Unit (IMU) and a 2D laser scanner, to concurrently estimate the six degree-of-freedom (d.o.f.) position and orientation (pose) of the sensing package and a 3D map of the environment. Rather than constraining the person to purely planar motion, the IMU measurements are integrated to estimate the pose along a general 3D trajectory. To mitigate the accumulation of inertial drift errors, the pose estimates are corrected using laser measurements, namely line-to-plane correspondences between linear segments in the laser-scan data and structural planes of the building. Utilizing orthogonal building planes as map features results in a human-interpretable layout of the environment, and ensures that the each feature can be efficiently initialized and estimated. A practical method is presented to initialize the pose and the IMU biases using observations of known planes and zero-velocity updates, respectively. In addition to the filter design, the observability properties of the nonlinear system are studied to show under which measurement conditions the 3D pose can be accurately estimated. Lastly, an approach for utilizing the sensors' measurements to perform on-line calibration of the laser-to-IMU transformation is developed, which enables the highest possible localization accuracy. The proposed L-INS is experimentally validated by a person traveling in both known and unknown 3D environments to demonstrate its reliability and accuracy for indoor localization and mapping.**

I. Introduction

For humans, safe and efficient navigation requires knowledge of the environmental layout, path planning, obstacle avoidance, and determining one's position and orientation (pose) with respect to the world. For a *visually-impaired* person, these tasks can be exceedingly difficult to accomplish, and there are high risks associated with failure in any of them. To address some of these issues, guide dogs and white canes are widely used for the purposes of wayfinding and environment sensing. The former, however, has costly training requirements, while the latter can only provide cues about one's immediate surroundings. On the other hand, commercially available electronic navigation systems designed for the visually impaired (e.g., Humanware (2010), Sendero (2010)) rely on GPS signals and cannot be utilized indoors, under tree cover, or next to tall buildings where reception is poor.

This work was supported by the University of Minnesota (DTC), the National Science Foundation (IIS-0643680, IIS-0811946, IIS-0835637), and AFOSR (FA9550-10-1-0567). J. A. Hesch was supported by the UMN Doctoral Dissertation Fellowship.

J. A. Hesch, F. M. Mirzaei, and S. I. Roumeliotis are with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA. G. L. Mariottini is with the Department of Computer Science and Engineering, University of Texas, Arlington, TX 76019, USA. E-mail: {joel|faraz|stergios}@cs.umn.edu, gianluca@uta.edu

Preliminary versions of this work appeared in Hesch et al. (2009) and Hesch et al. (2010).

In the academic community, numerous electronic navigation systems for GPS-denied environments have been proposed. However, the majority of the existing algorithms are designed for mobile robots that are limited to move on planar surfaces (Thrun et al. 2000; Iocchi and Pellegrini 2007) or require heavy sensors, such as a 3D laser scanner (Hähnel et al. 2003; Borrmann et al. 2008), that cannot be carried by a human. Other algorithms, which have relied on visual information (Kim and Sukkarieh 2007; Mourikis and Roumeliotis 2008), are sensitive to variable lighting conditions and require processing resources that are not typically available on portable computing devices.

To address these issues, we aim to design a personal *indoor* navigation system that fulfills the following requirements:

- The system must accurately track the *six degree-of-freedom (d.o.f.) pose* of the visually impaired person, allowing them to safely navigate in a *3D environment*.
- The navigation aid should enable the person to walk through previously *unknown buildings* without getting lost. This requires constructing a map of the explored area and localizing with respect to it in *real-time*.
- The selected sensors should be *robust* to environmental changes, such as lighting conditions, reliable in the presence of clutter and moving objects, and work within the *computational and memory limits* of a portable computing device.
- The navigation aid should be *compact*, *unobtrusive* to the person, and *lightweight* enough to be carried without fatigue.

To meet these objectives, we are focused on designing an indoor Laser-aided Inertial Navigation System (L-INS) using an *Inertial Measurement Unit (IMU)* and a *2D laser scanner*, based upon our preliminary results in Hesch et al. (2009, 2010). Employing this sensor pair ensures feasibility of manufacturing a light-weight and compact sensor package that can be carried by a person, since a wide variety of small IMUs (e.g., Memsense nIMU) and compact-size 2D laser scanners (e.g., Hokuyo URG) are commercially available. Additionally, using a laser scanner instead of a camera provides greater reliability and robustness under poor lighting conditions.

The proposed algorithm tracks the six d.o.f. pose of the person by integrating the IMU measurements (linear acceleration and rotational velocity) using an Extended Kalman Filter (EKF). However, without corrections from an exteroceptive sensor, the IMU measurement noise and bias drift would cause the pose estimation errors to grow unbounded over time. To mitigate this issue, we propose to update the pose estimates by utilizing straight-line features extracted from the 2D laser scans. In particular, as the person moves, the laser

scanner's attitude changes which causes its scanning plane to intersect a variety of structural planes of the building (i.e., the walls, floor, and ceiling). If the structural planes are known *a priori* from a building map, we can use the information from the line-to-plane measurements in order to update the person's pose estimates (Hesch et al. 2009). Unfortunately, in many cases in practice, a building map is not available in advance. To overcome this challenge, we simultaneously construct a building map in order to utilize *previously unknown* structural planes in the localization process (Hesch et al. 2010). We exploit the fact that most indoor structural planes are *orthogonal to each other*, which allows us to fix each plane's orientation the first time it is observed, and only estimate its distance to the origin of the global reference frame.

Constructing the map based on orthogonal planar structures has the advantage of keeping the person's orientation error bounded (Nguyen et al. 2006) in addition to providing inherent robustness to clutter and moving objects. Furthermore, the estimated map directly provides a *human-interpretable layout* of the building that can simplify the task of wayfinding towards a destination. Moreover, due to the limited number of structural planes in each building, the computational load of the algorithm remains bounded. This, together with the low processing cost of line-segment extraction from the 2D laser scans, ensures the real-time execution of the algorithm on a hand-held computer with limited computational and memory resources.

We demonstrate the validity and reliability of the proposed approach with real-world experiments in both known and unknown environments. In the first case, we present a loop trajectory of 120 m in length that covers part of one floor of the Keller Hall at the University of Minnesota. The second test covers multiple levels of Akerman Hall at the University of Minnesota. In this 270 m trajectory, the person traverses several staircases and a disability access ramp. In addition, both test environments includes significant clutter (e.g., trash-cans, storage boxes, and furniture), as well as a normal flow of pedestrian traffic. Despite these challenges, our algorithm accurately tracks the person's pose, and correctly estimates a map of the building layout.

In order to ensure that the IMU and the laser scanner measurements provide sufficient information for estimating the person's pose, we study the observability of the corresponding nonlinear system. We also address the more practical matter of how to efficiently initialize the filter. Lastly, we provide a novel on-line method for calibrating the laser-to-IMU transformation using either previously known or unknown planar features, since inaccurate calibration can lead to biased filter estimates.

The remainder of the paper is organized as follows: In Section II, we begin with an overview of the related literature. Section III presents the core of our algorithm, which is an EKF-based pose estimator. We describe how to efficiently initialize the state of the filter in Section IV. In Section V, we study the observability properties of the map-based localization system, and show the system is observable under mild conditions that are typically fulfilled in practice. Subsequently, we describe our approach for calibrating the laser-to-IMU transformation using line-to-plane correspondences in Section VI. Experimental validation of the proposed method is provided in Section VII. Lastly, we conclude the paper and present future research directions in Section VIII.

II. Related Work

Recent work has focused primarily on developing hazard-detection aids for the visually impaired with the purpose of *detecting and avoiding obstacles* (Ulrich and Borenstein 2001; Yuan and Manduchi 2005) and describing objects' size and color (Hub et al. 2004). These systems cannot be directly used as wayfinding aids without the development of appropriate algorithms for localization. In contrast to the above systems, navigation aids have been designed that explicitly track a person's location and heading direction. Most relevant efforts have primarily addressed GPS-based *outdoor navigation* which cannot be used inside a building (Makino et al. 1996; Ran et al. 2004). *Indoor navigation* is more challenging, since pose information can only be inferred from ego-motion and environmental cues. In what follows, we provide a discussion of several existing indoor navigation systems.

1) Navigating using ego-motion: Dead-reckoning systems track a person's pose *without any external references*. Common approaches are based on foot-mounted accelerometers (Cavallo et al. 2005). As a person walks, their position is computed by double integration of the acceleration measurements. Unfortunately, the accelerometer bias and noise are integrated as well, which causes the *position error* to grow unbounded. Even if the rate of position-error increase can be reduced with static-period drift corrections (Sagawa et al. 2000; Borenstein et al. 2009), dead-reckoning systems still remain unreliable over long time intervals.

2) Navigating with known references: Unlike dead-reckoning approaches that do not employ external references, map-based systems infer position and orientation information from known landmarks or beacons in the environment. For example, in Kulyukin et al. (2004), a robot is attached at the end of a leash as a substitute for a guide dog, and localizes using odometry and a network of Radio Frequency IDentification (RFID) tags. Tjan et al. (2005) presented another approach in which a hand-held camera identifies retro-reflective digital signs. Similar methods also exist based on ultrasound (Ran et al. 2004) and infrared (Ertan et al. 1998) beacons. In our previous work (Hesch and Roumeliotis 2010), we presented a map-based indoor localization aid for the visually impaired comprised of a pedometer, a tri-axial gyroscope, and a 2D laser scanner. We exploited known corners at hallway intersections (computed from the building blueprints) as landmarks for localization. Unfortunately, all map-based or beacon-based localization methods suffer from common limitations which include: (i) *time and cost* associated with acquiring the map or installing the beacons, (ii) the system's *inability to adapt* to spatial layout changes, and (iii) the *restriction* of use to previously mapped areas.

3) Navigating in unknown environments: The most flexible navigation aids are those that can exploit environment sensing to perform Simultaneous Localization and Mapping (SLAM). The majority of the proposed systems for SLAM

consider either 2D map and sensor motion (Smith et al. 1990; Dissanayake et al. 2001), or restrict the sensor motion to planar surfaces and create a 3D map of the surroundings (Thrun et al. 2000; Iocchi and Pellegrini 2007; Nguyen et al. 2006). These algorithms are not generally suitable for use on a personal navigation system since the motion of a human is not limited to a planar surface (e.g., when climbing stairs).

There exist several approaches for estimating a 3D map and the six d.o.f. pose of a robot (3D SLAM) that employ 3D point cloud matching techniques [e.g., Iterative Closest Point (ICP)] (Hähnel et al. 2003; Borrmann et al. 2008; Kohlhepp et al. 2004; Nüchter et al. 2004; Cole and Newman 2006). However, the computational requirements for matching 3D point clouds are typically prohibitive for real-time implementation. More importantly, the 3D laser scanners needed for acquiring the point clouds are too large and heavy for a person to carry, thus making these systems inappropriate for use as a personal navigation aid. Alternative methods for performing 3D SLAM employ cameras to map the environment based on visual landmarks (Mourikis and Roumeliotis 2008; Kim and Sukkarieh 2007). The main drawback of camera-based systems is their sensitivity to variable lighting conditions, which restricts their use as navigation aids for the visually impaired where reliability is of paramount importance. Additionally, processing images and extracting visual features are typically computationally intensive tasks that are impractical to carry out on hand-held computing devices. Furthermore, constructing a map of the 3D locations of visual landmarks (e.g., SIFT features (Lowe 2004)) often used in these approaches may not be geometrically meaningful or interpretable by humans. Finally, extracting and matching visual landmarks in indoor environments can be challenging and unreliable due to insufficient texture.

To address these limitations, we propose an L-INS based on a 2D laser scanner and an IMU. The key differentiating factor of our work is that we can explore and map 3D environments with a sensing package that follows arbitrary 3D trajectories, despite the fact that the exteroceptive sensor employed only senses in 2D during each laser scan. Specifically, our system tracks the six d.o.f. pose of the person and measures both known building planes as well as new planes which it maps as the unknown portions of the environment are explored. We note that using commonly-occurring structural planes as map features ensures the applicability of the method in practice. The estimated structural planes directly represent the geometric layout of the building that can be easily interpreted by humans. Moreover, due to the limited number of structural planes in each building, the computational requirements of our algorithm do not grow unbounded over time, since the size of the estimated state vector remains bounded. Finally, our algorithm can perform *on-line calibration* of the relative pose between laser and IMU, which was assumed known in our previous contributions (Hesch et al. 2009; Hesch and Roumeliotis 2010).

III. Algorithm Description

A hand-held computer collects measurements from the navigation aid consisting of an IMU and a 2D laser scanner, which are rigidly connected (see Fig. 1). The sensor data is fused in an EKF to concurrently estimate the six d.o.f. pose of the sensor platform, as well as the 3D map of the building's perpendicular structural planes (i.e., the walls, floor, and ceiling). In what follows, we present the propagation and update models used by the EKF.

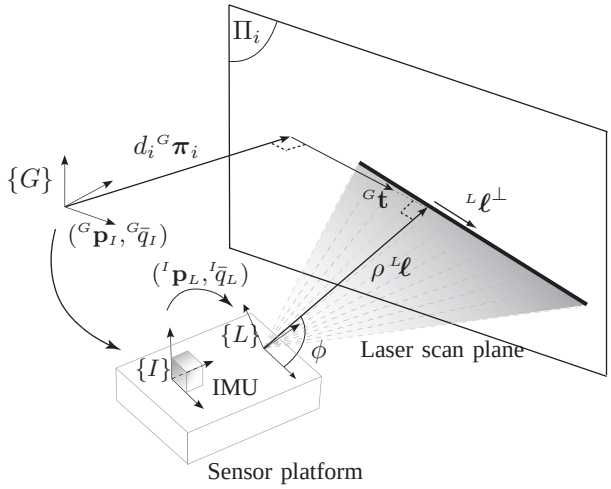


Fig. 1. As the IMU-laser sensor platform moves, the laser scan plane intersects a structural planar surface, Π_i, described by d_i and ${}^G\boldsymbol{\pi}_i$, which are the Hessian normal form components of the plane with respect to the global frame of reference, $\{G\}$. The shortest vector in the laser scan plane from the origin of the laser frame, $\{L\}$, to Π_i has length ρ and direction ${}^L\boldsymbol{\ell}$, with respect to $\{L\}$. The line of intersection has direction, ${}^L\boldsymbol{\ell}^\perp$, with respect to $\{L\}$ and is described by the polar parameters (ρ, ϕ). The vector from the intersection of ${}^G\boldsymbol{\pi}_i$ and Π_i to the intersection of $\rho\,{}^L\boldsymbol{\ell}$ and Π_i, is ${}^G\mathbf{t}$. The IMU-laser transformation is denoted by $({}^I\mathbf{p}_L, {}^I\bar{q}_L)$, while the IMU pose with respect to $\{G\}$ is $({}^G\mathbf{p}_I, {}^G\bar{q}_I)$.

A. Filter Propagation

The EKF estimates the IMU pose and linear velocity together with the time-varying IMU biases and the map. The filter state is the $(16+N)\times 1$ vector:

$$\begin{aligned}\mathbf{x} &= \begin{bmatrix} {}^I\bar{q}_G^T & \mathbf{b}_g^T & {}^G\mathbf{v}_I^T & \mathbf{b}_a^T & {}^G\mathbf{p}_I^T & | & d_1 \cdots d_N \end{bmatrix}^T \\ &= \begin{bmatrix} \mathbf{x}_s^T & | & \mathbf{x}_d^T \end{bmatrix}^T, \end{aligned} \tag{1}$$

where $\mathbf{x}_s(t)$ is the 16×1 sensor platform state, and $\mathbf{x}_d(t)$ is the $N\times 1$ state of the structural plane map. The first component of the sensor platform state is, ${}^I\bar{q}_G(t)$, which is the unit quaternion representing the orientation of the *global frame* $\{G\}$ in the IMU frame, $\{I\}$, at time t. The frame $\{I\}$ is attached to the IMU (see Fig. 1), while $\{G\}$ is an inertial reference frame whose origin coincides with the initial IMU position, and whose orientation is aligned with the perpendicular structural planes according to the filter initialization procedure described in Section IV. The sensor platform state also includes the position and velocity of $\{I\}$ in $\{G\}$, denoted by the 3×1 vectors ${}^G\mathbf{p}_I(t)$ and ${}^G\mathbf{v}_I(t)$, respectively. The remaining components are the biases, $\mathbf{b}_g(t)$ and $\mathbf{b}_a(t)$, affecting the gyroscope and accelerometer measurements, which are modeled as random-walk processes driven by the zero-mean, white Gaussian noise $\mathbf{n}_{wg}(t)$ and $\mathbf{n}_{wa}(t)$, respectively.

The building map is comprised of N static planar features Π_i, $i = 1, \ldots, N$, which includes all planes (if any) that are known from the building blue prints, and grows as new planes are detected. Each plane is described by its Hessian normal form components d_i and ${}^G\boldsymbol{\pi}_i$, which are the distance from the plane to the origin of $\{G\}$, and the 3×1 normal vector of the plane expressed in $\{G\}$, respectively.[1] The map state, $\mathbf{x}_d$, consists of the scalar distances, d_i, $i = 1, \ldots, N$, which are estimated along with the state of the sensing package. We only map perpendicular structural planes, hence, we do not need to estimate each plane's normal-vector. Instead, we store them in the map parameter vector $\left[{}^G\boldsymbol{\pi}_1^T \ldots {}^G\boldsymbol{\pi}_N^T\right]^T$, where each component ${}^G\boldsymbol{\pi}_i$ is determined once during the new plane initialization step (see Section III-C) or is available from the blueprint layout. With the state of the system now defined, we turn our attention to the continuous-time dynamical model which governs the state of the system.

1) Continuous-time model: The system model describing the time evolution of the state is (see Lefferts et al. (1982); Trawny and Roumeliotis (2005)):

$$
\begin{aligned}
{}^I\dot{\bar{q}}_G(t) &= \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega}(t)){}^I\bar{q}_G(t) && (2)\\
{}^G\dot{\mathbf{p}}_I(t) &= {}^G\mathbf{v}_I(t) \;\;,\;\; {}^G\dot{\mathbf{v}}_I(t) = {}^G\mathbf{a}(t) && (3)\\
\dot{\mathbf{b}}_g(t) &= \mathbf{n}_{wg}(t) \;\;,\;\; \dot{\mathbf{b}}_a(t) = \mathbf{n}_{wa}(t) && (4)\\
\dot{d}_i(t) &= 0 \;,\; i = 1, \ldots, N. && (5)
\end{aligned}
$$

In these expressions, $\boldsymbol{\omega}(t) = [\omega_1(t) \;\; \omega_2(t) \;\; \omega_3(t)]^T$ is the rotational velocity of the IMU, expressed in $\{I\}$, ${}^G\mathbf{a}$ is the IMU acceleration expressed in $\{G\}$, and

$$
\boldsymbol{\Omega}(\boldsymbol{\omega}) = \begin{bmatrix} -\lfloor\boldsymbol{\omega}\times\rfloor & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^T & 0 \end{bmatrix}, \quad \lfloor\boldsymbol{\omega}\times\rfloor \triangleq \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}.
$$

The gyroscope and accelerometer measurements, $\boldsymbol{\omega}_m$ and $\mathbf{a}_m$, used for state propagation, are

$$
\begin{aligned}
\boldsymbol{\omega}_m(t) &= \boldsymbol{\omega}(t) + \mathbf{b}_g(t) + \mathbf{n}_g(t) && (6)\\
\mathbf{a}_m(t) &= \mathbf{C}({}^I\bar{q}_G(t))\,({}^G\mathbf{a}(t) - {}^G\mathbf{g}) + \mathbf{b}_a(t) + \mathbf{n}_a(t), && (7)
\end{aligned}
$$

where $\mathbf{n}_g$ and $\mathbf{n}_a$ are zero-mean, white Gaussian noise processes, and ${}^G\mathbf{g}$ is the gravitational acceleration. The matrix $\mathbf{C}(\bar{q})$ is the rotation matrix corresponding to $\bar{q}$. Also note that the distances to the building planes are fixed with respect to $\{G\}$, thus their time derivatives are zero [see (5)].

Linearizing at the current estimates and applying the expectation operator on both sides of (2)-(5), we obtain the state estimate propagation model

$$
\begin{aligned}
{}^I\dot{\hat{\bar{q}}}_G(t) &= \frac{1}{2}\boldsymbol{\Omega}(\hat{\boldsymbol{\omega}}(t)){}^I\hat{\bar{q}}_G(t) && (8)\\
{}^G\dot{\hat{\mathbf{p}}}_I(t) &= {}^G\hat{\mathbf{v}}_I(t) \;\;,\;\; {}^G\dot{\hat{\mathbf{v}}}_I(t) = \mathbf{C}^T({}^I\hat{\bar{q}}_G(t))\,\hat{\mathbf{a}}(t) + {}^G\mathbf{g} && (9)\\
\dot{\hat{\mathbf{b}}}_g(t) &= \mathbf{0}_{3\times1} \;\;,\;\; \dot{\hat{\mathbf{b}}}_a(t) = \mathbf{0}_{3\times1} && (10)\\
\dot{\hat{d}}_i\,(t) &= 0 \;\;,\;\; i = 1, \ldots, N, && (11)
\end{aligned}
$$

with $\hat{\mathbf{a}}(t) = \mathbf{a}_m(t) - \hat{\mathbf{b}}_a(t)$, and $\hat{\boldsymbol{\omega}}(t) = \boldsymbol{\omega}_m(t) - \hat{\mathbf{b}}_g(t)$.

[1] A point ${}^G\mathbf{p}$ lies on plane Π_i if ${}^G\boldsymbol{\pi}_i^T{}^G\mathbf{p} - d_i = 0$.

The $(15 + N) \times 1$ error-state vector is defined as

$$
\begin{aligned}
\widetilde{\mathbf{x}} &= \begin{bmatrix} {}^I\boldsymbol{\delta\theta}_G^T & \widetilde{\mathbf{b}}_g^T & {}^G\widetilde{\mathbf{v}}_I^T & \widetilde{\mathbf{b}}_a^T & {}^G\widetilde{\mathbf{p}}_I^T & | & \widetilde{d}_1 \cdots \widetilde{d}_N \end{bmatrix}^T \\
&= \begin{bmatrix} \widetilde{\mathbf{x}}_s^T & | & \widetilde{\mathbf{x}}_d^T \end{bmatrix}^T, && (12)
\end{aligned}
$$

where $\widetilde{\mathbf{x}}_s(t)$ is the 15×1 error state corresponding to the sensing platform, and $\widetilde{\mathbf{x}}_d(t)$ is the $N \times 1$ error state of the map. For the IMU position, velocity, biases, and the map, an additive error model is utilized (i.e., $\widetilde{x} = x - \hat{x}$ is the error in the estimate $\hat{x}$ of a quantity x). However, for the quaternion we employ a multiplicative error model. Specifically, the error between the quaternion $\bar{q}$ and its estimate $\hat{\bar{q}}$ is the 3×1 angle-error vector, $\boldsymbol{\delta\theta}$, implicitly defined by the *error quaternion*

$$
\delta\bar{q} = \bar{q} \otimes \hat{\bar{q}}^{-1} \simeq \begin{bmatrix} \frac{1}{2}\boldsymbol{\delta\theta}^T & 1 \end{bmatrix}^T, \qquad (13)
$$

where $\delta\bar{q}$ describes the small rotation that causes the true and estimated attitude to coincide. The main advantage of this error definition is that it allows us to represent the attitude uncertainty by the 3×3 covariance matrix $E\{\boldsymbol{\delta\theta}\boldsymbol{\delta\theta}^T\}$. Since the attitude corresponds to three d.o.f., this is a minimal representation.

The linearized continuous-time error-state equation is

$$
\begin{aligned}
\dot{\widetilde{\mathbf{x}}} &= \begin{bmatrix} \mathbf{F}_{s,c} & \mathbf{0}_{15\times N} \\ \mathbf{0}_{N\times15} & \mathbf{I}_N \end{bmatrix} \widetilde{\mathbf{x}} + \begin{bmatrix} \mathbf{G}_{s,c} \\ \mathbf{0}_{N\times15} \end{bmatrix} \mathbf{n} \\
&= \mathbf{F}_c\,\widetilde{\mathbf{x}} + \mathbf{G}_c\,\mathbf{n}\,, && (14)
\end{aligned}
$$

where $\mathbf{I}_N$ denotes the $N \times N$ identity matrix, $\mathbf{F}_{s,c}$ is the continuous-time error-state transition matrix corresponding to the sensor platform state, and $\mathbf{G}_{s,c}$ is the continuous time input noise matrix, i.e.,

$$
\mathbf{F}_{s,c} = \begin{bmatrix} -\lfloor\hat{\boldsymbol{\omega}}\times\rfloor & -\mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ -\mathbf{C}^T({}^I\hat{\bar{q}}_G)\lfloor\hat{\mathbf{a}}\times\rfloor & \mathbf{0}_3 & \mathbf{0}_3 & -\mathbf{C}^T({}^I\hat{\bar{q}}_G) & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}
$$

$$
\mathbf{G}_{s,c} = \begin{bmatrix} -\mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & -\mathbf{C}^T({}^I\hat{\bar{q}}_G) & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \end{bmatrix}, \; \mathbf{n} = \begin{bmatrix} \mathbf{n}_g \\ \mathbf{n}_{wg} \\ \mathbf{n}_a \\ \mathbf{n}_{wa} \end{bmatrix},
$$

where $\mathbf{0}_3$ is the 3×3 matrix of zeros. The system noise covariance matrix $\mathbf{Q}_c$ depends on the IMU noise characteristics and is computed off-line (Trawny and Roumeliotis 2005).

2) Discrete-time implementation: The IMU signals $\boldsymbol{\omega}_m$ and $\mathbf{a}_m$ are sampled at a constant rate $1/T$, where $T \triangleq t_{k+1} - t_k$. Every time a new IMU measurement is received, the state estimate is propagated using 4th-order Runge-Kutta numerical integration of (8)–(11). In order to derive the discrete-time covariance propagation equation, we evaluate the discrete-time state transition matrix

$$
\boldsymbol{\Phi}_k = \boldsymbol{\Phi}(t_{k+1}, t_k) = \exp\left(\int_{t_k}^{t_{k+1}} \mathbf{F}_c(\tau)\mathrm{d}\tau\right) \qquad (15)
$$

and the discrete-time system noise covariance matrix

$$
\mathbf{Q}_{d,k} = \int_{t_k}^{t_{k+1}} \boldsymbol{\Phi}(t_{k+1}, \tau)\mathbf{G}_c\mathbf{Q}_c\mathbf{G}_c^T\boldsymbol{\Phi}^T(t_{k+1}, \tau)\mathrm{d}\tau. \qquad (16)
$$

The propagated covariance is then computed as

$$\mathbf{P}_{k+1|k} = \mathbf{\Phi}_k \mathbf{P}_{k|k} \mathbf{\Phi}_k^T + \mathbf{Q}_{d,k}. \quad (17)$$

After processing the IMU measurements to propagate the filter state and covariance, we process any available laser scan measurements in the filter update step (see Section III-B).

B. Landmark Update

As the IMU-laser platform moves in an indoor environment, the laser-scan plane intersects the perpendicular structural planes of the building. These measurements are exploited to update the state estimate. To simplify the discussion, we consider a single line measurement, ${}^L\boldsymbol{\ell}^\perp$, corresponding to the intersection of the laser-scan plane and map plane, Π_i (see Fig. 1). The line is described in the laser frame, $\{L\}$, by (ρ, ϕ), where ρ is the distance from the origin of $\{L\}$ to the line, and ϕ is the angle of the vector ${}^L\boldsymbol{\ell}$ perpendicular to the line.[2] We will hereafter express the line direction in $\{I\}$, as ${}^I\boldsymbol{\ell}^\perp = \mathbf{C}({}^I\bar{q}_L)\begin{bmatrix}\sin\phi & -\cos\phi & 0\end{bmatrix}^T$, where ${}^I\bar{q}_L$ is the unit quaternion representing the orientation of the laser frame in the IMU frame (see Sect. VI). In what follows, we describe how each line is exploited to define two *measurement constraints*, which are used by the EKF to update the state estimates.

1) Orientation Constraint: The first constraint is on the orientation of $\{I\}$ with respect to $\{G\}$. The normal to the plane Π_i, vector ${}^G\boldsymbol{\pi}_i$, is perpendicular to $\mathbf{C}({}^I\bar{q}_G)\,{}^I\boldsymbol{\ell}^\perp$ (see Fig. 1), which yields the following *orientation measurement constraint*

$$z_1 = {}^G\boldsymbol{\pi}_i^T\,\mathbf{C}({}^I\bar{q}_G)\,{}^I\boldsymbol{\ell}^\perp = 0\,. \quad (18)$$

The expected measurement is

$$\hat{z}_1 = {}^G\boldsymbol{\pi}_i^T\,\mathbf{C}({}^I\hat{\bar{q}}_G)\,{}^I\boldsymbol{\ell}_m^\perp\,, \quad (19)$$

where ${}^I\boldsymbol{\ell}_m^\perp = \mathbf{C}({}^I\bar{q}_L)\begin{bmatrix}\sin\phi_m & -\cos\phi_m & 0\end{bmatrix}^T$ is the *measured* line direction with $\phi_m = \phi - \tilde{\phi}$. The measurement residual is $r_1 = z_1 - \hat{z}_1 = -\hat{z}_1$ and the corresponding linearized error model is

$$\begin{aligned}\widetilde{z}_1 \simeq\ & \begin{bmatrix}-{}^G\boldsymbol{\pi}_i^T\mathbf{C}({}^I\hat{\bar{q}}_G)\lfloor {}^I\boldsymbol{\ell}_m^\perp\times\rfloor & \mathbf{0}_{1\times 12}\end{bmatrix}\widetilde{\mathbf{x}}_s \\ & + \begin{bmatrix}\mathbf{0}_{1\times N}\end{bmatrix}\widetilde{\mathbf{x}}_d + \begin{bmatrix}{}^G\boldsymbol{\pi}_i^T\mathbf{C}({}^I\hat{\bar{q}}_G){}^I\boldsymbol{\ell}_m & 0\end{bmatrix}\begin{bmatrix}\tilde{\phi}\\ \tilde{\rho}\end{bmatrix} \\ =\ & \mathbf{h}_{1,s}^T\,\widetilde{\mathbf{x}}_s + \mathbf{h}_{1,d}^T\,\widetilde{\mathbf{x}}_d + \boldsymbol{\gamma}_1^T\,\mathbf{n}_\ell, \end{aligned} \quad (20)$$

where ${}^I\boldsymbol{\ell}_m = \mathbf{C}({}^I\bar{q}_L)\begin{bmatrix}\cos\phi_m & \sin\phi_m & 0\end{bmatrix}^T$ is the perpendicular to the measured line direction and $\rho_m = \rho - \tilde{\rho}$ is the measured distance to the line. The vectors $\mathbf{h}_{1,s}^T$, $\mathbf{h}_{1,d}^T$, and $\boldsymbol{\gamma}_1^T$ are the Jacobians of (18) with respect to the state and line parameters, respectively. The 2×1 error vector $\mathbf{n}_\ell$ is assumed to be zero-mean, white Gaussian, with covariance matrix $\mathbf{R} = E\{\mathbf{n}_\ell\mathbf{n}_\ell^T\}$ computed for each line from the weighted line-fitting procedure (Pfister et al. 2003).

2) Distance Constraint: From Fig. 1, the following geometric relationship holds:

$${}^G\mathbf{p}_I + \mathbf{C}({}^I\bar{q}_G)\,({}^I\mathbf{p}_L + \rho\,{}^I\boldsymbol{\ell}) = d_i\,{}^G\boldsymbol{\pi}_i + {}^G\mathbf{t}, \quad (21)$$

[2]We utilized the Split-and-Merge algorithm (Nguyen et al. 2007) to segment the laser-scan data and a weighted line-fitting algorithm (Pfister et al. 2003) to estimate the line parameters (ρ, ϕ) for each line.

where ${}^I\boldsymbol{\ell} = \mathbf{C}({}^I\bar{q}_L)\begin{bmatrix}\cos\phi & \sin\phi & 0\end{bmatrix}^T$ is the perpendicular to the line direction, and ${}^I\mathbf{p}_L$ is the position of the laser scanner in the IMU frame. Since the vector ${}^G\mathbf{t}$ is unknown and cannot be measured we need to eliminate it from the equation. We do so by projecting (21) onto ${}^G\boldsymbol{\pi}_i^T$, yielding the *distance measurement constraint*

$$z_2 = {}^G\boldsymbol{\pi}_i^T\,({}^G\mathbf{p}_I + \mathbf{C}({}^I\bar{q}_G)\,({}^I\mathbf{p}_L + \rho\,{}^I\boldsymbol{\ell})) - d_i = 0. \quad (22)$$

The expected measurement is

$$\hat{z}_2 = {}^G\boldsymbol{\pi}_i^T\,\left({}^G\hat{\mathbf{p}}_I + \mathbf{C}({}^I\hat{\bar{q}}_G)\,({}^I\mathbf{p}_L + \rho_m\,{}^I\boldsymbol{\ell}_m)\right) - \hat{d}_i. \quad (23)$$

The measurement residual is $r_2 = z_2 - \hat{z}_2 = -\hat{z}_2$ and the corresponding linearized error model is

$$\begin{aligned}\widetilde{z}_2 \simeq\ & \begin{bmatrix}-{}^G\boldsymbol{\pi}_i^T\mathbf{C}^T\left({}^I\hat{\bar{q}}_G\right)\lfloor {}^I\mathbf{p}_L + \rho_m\,{}^I\boldsymbol{\ell}_m\times\rfloor & \mathbf{0}_{1\times 9} & {}^G\boldsymbol{\pi}_i^T\end{bmatrix}\widetilde{\mathbf{x}}_s \\ & + \begin{bmatrix}\mathbf{0}_{1\times(i-1)} & -1 & \mathbf{0}_{1\times(N-i)}\end{bmatrix}\widetilde{\mathbf{x}}_d \\ & + \begin{bmatrix}-{}^G\boldsymbol{\pi}_i^T\mathbf{C}^T\left({}^I\hat{\bar{q}}_G\right)\rho_m\,{}^I\boldsymbol{\ell}_m^\perp & {}^G\boldsymbol{\pi}_i^T\mathbf{C}^T\left({}^I\hat{\bar{q}}_G\right){}^I\boldsymbol{\ell}_m\end{bmatrix}\begin{bmatrix}\tilde{\phi}\\ \tilde{\rho}\end{bmatrix} \\ =\ & \mathbf{h}_{2,s}^T\,\widetilde{\mathbf{x}}_s + \mathbf{h}_{2,d}^T\,\widetilde{\mathbf{x}}_d + \boldsymbol{\gamma}_2^T\,\mathbf{n}_\ell, \end{aligned} \quad (24)$$

where the vectors $\mathbf{h}_{2,s}^T$, $\mathbf{h}_{2,d}^T$, and $\boldsymbol{\gamma}_2^T$ are the Jacobians of (22) with respect to the state and line parameters, respectively.

We process the two measurement constraints together; stacking (20) and (24), we obtain the measurement Jacobians

$$\mathbf{H} = \begin{bmatrix}\mathbf{h}_{1,s}^T & \mathbf{h}_{1,d}^T \\ \mathbf{h}_{2,s}^T & \mathbf{h}_{2,d}^T\end{bmatrix}\quad,\quad \mathbf{\Gamma} = \begin{bmatrix}\boldsymbol{\gamma}_1^T \\ \boldsymbol{\gamma}_2^T\end{bmatrix}, \quad (25)$$

which are used in the expression for the Kalman gain

$$\mathbf{K} = \mathbf{P}_{k+1|k}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^T + \mathbf{\Gamma}\mathbf{R}\mathbf{\Gamma}^T\right)^{-1}. \quad (26)$$

The residual vector is $\mathbf{r} = \begin{bmatrix}r_1 & r_2\end{bmatrix}^T$, and the state and the covariance update equations are

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}\mathbf{r}$$

$$\mathbf{P}_{k+1|k+1} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}_{k+1|k}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{\Gamma}\mathbf{R}\mathbf{\Gamma}^T\mathbf{K}^T.$$

After updating the state and covariance with measurements to planes currently in the map, we may have additional measurements to process corresponding to planes that have not been observed previously. In Section III-C we describe how to augment the map with an initial estimate of each new feature.

C. Landmark Initialization

There are three cases which we distinguish for plane initialization. The first is planes which are known perfectly *a priori* (e.g., from "as-built" building blueprints). The second class are planes which are approximately known (e.g., extracted from imprecise building blueprints, or "as-designed"). The third type are the unknown planes that occur in the principal building directions (i.e., the floor, ceiling, and orthogonal building walls). While we do not know the location or number of these planes, whenever we observe them, we know they exhibit one of the three known principle orientations, and only the distance to the plane must be estimated.

1) Perfectly known planes: Perfectly known planes are straight forward to exploit in our navigation framework since

all three d.o.f. of the plane parameters are known *a priori*. We could include each plane in the state vector with an associated zero-covariance and zero-correlation to the rest of the state. However, in practice we simply maintain an additional parameter vector of known planes, which reduces the computational cost of the filter by limiting the state size even further. When observing a perfectly known plane, we follow the procedure in Sect. III-B to update the state, with the caveat that the Jacobians taken with respect to the known plane parameters are set identically to zero [see (20) and (24)].

2) Approximately known planes: Planes which are known approximately are the most common to arise in typical implementations when a blueprint of the building is available. This occurs because for practical reasons during building construction, walls are not always placed precisely where they were designated and building tolerances permit some room for error. In these instances, we include an initial estimate of each building plane in the map, and we set the covariance for each plane according to the accuracy of the blueprints. In practice, if the quality of the blueprints is unknown, it suffices to hand-measure a small number of the building planes in order to characterize the blueprint accuracy. We assume that the errors in the initial estimates of the approximately known planes are uncorrelated with each other and the sensor platform state, and set the corresponding cross-correlation entries in $\mathbf{P}$ to zero.

3) Unknown planes: When measuring a new plane, Π_{N+1}, we first determine if the plane's orientation, ${}^G\boldsymbol{\pi}_{N+1}$, corresponds to one of the three cardinal directions, $\mathbf{e}_j$, $j=1,2,3$, considered in the map. We employ a Mahalanobis distance test to measure the probability of correspondence between the plane's orientation and each of the cardinal directions in the map. Specifically, we compute the orientation residual $r_{1,j} = -\mathbf{e}_j^T \mathbf{C}^T\big({}^I\hat{\bar{q}}_G\big)\,{}^I\boldsymbol{\ell}_m^{\perp}$, $j=1,2,3$, and the covariance of the residual

$$s_j = \begin{bmatrix}\mathbf{h}_{1,s}^T & \mathbf{h}_{1,d}^T\end{bmatrix}\mathbf{P}_{k+1|k}\begin{bmatrix}\mathbf{h}_{1,s}\\ \mathbf{h}_{1,d}\end{bmatrix} + \sigma_\phi^2 \boldsymbol{\gamma}_1^T\boldsymbol{\gamma}_1, \qquad (27)$$

where $\mathbf{h}_{1,s}$ and $\boldsymbol{\gamma}_1$ are the measurement Jacobians defined in (20) evaluated at ${}^G\boldsymbol{\pi}_i = \mathbf{e}_j$, and σ_ϕ^2 is the $(1,1)$ element of $\mathbf{R}$. If the smallest Mahalanobis distance

$$\mu_{jmin}^2 = \min_j \frac{r_{1,j}^2}{s_j} \qquad (28)$$

is less than a probabilistic threshold, then a new landmark is initialized with normal vector ${}^G\boldsymbol{\pi}_{N+1} = \mathbf{e}_{jmin}$. After determining the new plane's orientation, we compute the distance to the new plane by solving (23) for $\hat{d}_{N+1}$, i.e.,

$$\hat{d}_{N+1} = {}^G\boldsymbol{\pi}_{N+1}^T\big({}^G\hat{\mathbf{p}}_I + \mathbf{C}^T({}^I\hat{\bar{q}}_G)\,({}^I\mathbf{p}_L + \rho_m\,{}^I\boldsymbol{\ell}_m)\big) \qquad (29)$$

and augment the state vector as $\hat{\mathbf{x}}^{aug} \triangleq \begin{bmatrix}\hat{\mathbf{x}}^T & | & \hat{d}_{N+1}\end{bmatrix}^T$. Next, we need to augment the filter's covariance, which requires first partitioning the prior covariance into

$$\mathbf{P}_{k+1|k} = \begin{bmatrix}\mathbf{P}_{ss} & \mathbf{P}_{sd}\\ \mathbf{P}_{ds} & \mathbf{P}_{dd}\end{bmatrix}, \qquad (30)$$

where $\mathbf{P}_{ss}$ is the 15×15 sensor error-state covariance, $\mathbf{P}_{dd}$ is the $N\times N$ map error-state covariance, and $\mathbf{P}_{sd} = \mathbf{P}_{ds}^T$ are the $15\times N$ cross-correlation components. We then compute the scalar variance of the new plane, $\mathbf{P}_{d'd'}$, and the correlation between the new plane and the current state, $\mathbf{P}_{d'\mathbf{x}}$, as:

$$\mathbf{P}_{d'd'} = \mathbf{h}_{2,s}^T\mathbf{P}_{ss}\mathbf{h}_{2,s} + \boldsymbol{\gamma}_2^T\mathbf{R}\boldsymbol{\gamma}_2 \qquad (31)$$

$$\mathbf{P}_{d'\mathbf{x}} = \mathbf{P}_{\mathbf{x}d'}^T = \begin{bmatrix}\mathbf{h}_{2,s}^T\mathbf{P}_{ss} & \mathbf{h}_{2,s}^T\mathbf{P}_{sd}\end{bmatrix} \qquad (32)$$

where $\mathbf{h}_{2,s}$ and $\boldsymbol{\gamma}_2$ are defined in (24). Lastly, the augmented covariance, $\mathbf{P}^{aug}$, is computed as:

$$\mathbf{P}^{aug} = \begin{bmatrix}\mathbf{P}_{k+1|k} & \mathbf{P}_{\mathbf{x}d'}\\ \mathbf{P}_{d'\mathbf{x}} & \mathbf{P}_{d'd'}\end{bmatrix}. \qquad (33)$$

After performing state and covariance augmentation during the landmark initialization step, we return to the propagation step and process the next IMU measurement (see Section III-A).

D. Zero-Velocity Update

When the laser scanner does not detect any structural planes along certain directions for an extended period of time, the pose estimates accumulate errors due to drifts in the accelerometer and gyroscope biases. In addition, build up of orientation errors can cause the filter to incorrectly integrate a portion of the gravitational acceleration. This effect is closely related to the system's observability (see Section V) and can be compensated by means of drift correction during instantaneous stationary periods of the motion (e.g., when a shoe-mounted IMU is stationary during the stance phase while walking, see Sagawa et al. (2000)).

This procedure, termed a *zero-velocity update*, is challenging for two reasons: (i) the stationary periods must be identified without an *external reference*, and (ii) the IMU drift error must be corrected while properly accounting for the state uncertainty and IMU noise. Existing methods typically detect stationary periods based on a threshold check for the accelerometer measurement. These require significant hand tuning, and cannot account for the uncertainty in the current state estimate.

In contrast, we formulate the zero-velocity constraint as an EKF measurement and use the Mahalanobis distance test to identify the stationary intervals. Specifically, for the zero-velocity update, we employ the following measurement constraints for the linear acceleration, and linear and rotational velocities which are (instantaneously) equal to zero

$$\mathbf{z}_\zeta = \begin{bmatrix}\mathbf{a}^T & \boldsymbol{\omega}^T & {}^G\mathbf{v}_I^T\end{bmatrix}^T = \mathbf{0}_{9\times 1}. \qquad (34)$$

The zero-velocity measurement residual is

$$\mathbf{r}_\zeta = \mathbf{z}_\zeta - \hat{\mathbf{z}}_\zeta = \begin{bmatrix}\mathbf{a}_m - \hat{\mathbf{b}}_a + \mathbf{C}\big({}^I\hat{\bar{q}}_G\big){}^G\mathbf{g}\\ \boldsymbol{\omega}_m - \hat{\mathbf{b}}_g\\ -{}^G\hat{\mathbf{v}}_I\end{bmatrix} \qquad (35)$$

and the corresponding linearized error model is

$$\begin{aligned}\tilde{\mathbf{z}}_\zeta &\simeq \begin{bmatrix}-\lfloor \mathbf{C}\big({}^I\hat{\bar{q}}_G\big){}^G\mathbf{g}\times\rfloor & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} & \mathbf{I}_3 & \mathbf{0}_{3\times 3}\\ \mathbf{0}_{3\times 3} & \mathbf{I}_3 & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3}\\ \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3} & \mathbf{I}_3 & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 3}\end{bmatrix}\tilde{\mathbf{x}} + \begin{bmatrix}\mathbf{n}_a\\ \mathbf{n}_g\\ \mathbf{n}_v\end{bmatrix}\\ &= \mathbf{H}_\zeta\,\tilde{\mathbf{x}} + \mathbf{n}_\zeta\,, \end{aligned} \qquad (36)$$

where $\mathbf{H}_\zeta$ is the Jacobian of the zero-velocity measurement with respect to the state, and $\mathbf{n}_v$ is a zero-mean, white Gaussian process noise that acts as a regularization term

for computing the inverse of the measurement residuals' covariance. Based on this update model, at time step k we compute the Mahalanobis distance $\chi^2 = \mathbf{r}_\zeta^T \mathbf{S}_k^{-1} \mathbf{r}_\zeta$, where $\mathbf{S}_k = \mathbf{H}_\zeta \mathbf{P}_{k|k} \mathbf{H}_\zeta^T + \mathbf{R}_\zeta$ is the covariance of the measurement residual and $\mathbf{R}_\zeta = E\{\mathbf{n}_\zeta \mathbf{n}_\zeta^T\}$ is the measurement noise covariance. If χ^2 is less than a chosen probabilistic threshold, a stationary interval is detected and the state vector and the covariance matrix are updated using (34)-(36). We note that once we use the inertial measurements for an update, we cannot use them for propagation. However, this is not an issue, since the IMU is static and we do not need to use the kinematic model (2)-(5) to propagate the state estimates. Instead we employ the following equations:

$$ {}^I\dot{\bar{q}}_G(t) = \mathbf{0}_{4\times 1} \quad , \quad {}^G\dot{\mathbf{p}}_I(t) = \mathbf{0}_{3\times 1} \quad , \quad {}^G\dot{\mathbf{v}}_I(t) = \mathbf{0}_{3\times 1} $$
$$ \dot{\mathbf{b}}_g(t) = \mathbf{n}_{wg}(t) \quad , \quad \dot{\mathbf{b}}_a(t) = \mathbf{n}_{wa}(t). $$

In essence, this static-IMU propagation model indicates that the state vector and the covariance matrix of all components are kept constant. The only exceptions are the covariances of the errors in the gyroscope and accelerometer bias estimates which increase at each time step to reflect the effect of the random walk model.

IV. Filter State Initialization

Before using the EKF to fuse measurements from the laser scanner and the IMU, we need to initialize the state vector estimate $\widehat{\mathbf{x}}_{0|0}$ along with its covariance $\mathbf{P}_{0|0}$. This is performed in three sequential stages: (i) the gyroscopes' biases, $\mathbf{b}_g$, are initialized using the *partial* zero-velocity updates (Section IV-A), (ii) the IMU orientation, ${}^I\bar{q}_G$, is initialized employing the laser scans (Section IV-B), and (iii) the accelerometers' biases, $\mathbf{b}_a$, are initialized using zero-velocity updates (Section IV-C). Once these three stages are completed, we initialize the position of the sensing platform, ${}^G\mathbf{p}_I$. However, we note that if there are no structural planes in the building map initially known (i.e., if no blue print was provided), we can arbitrarily select the origin of the global frame. Thus, for our convenience, we set the origin of the global frame to coincide with the origin of the initial IMU frame, i.e., ${}^G\mathbf{p}_I = \mathbf{0}_{3\times 1}$. The initial covariance for ${}^G\widetilde{\mathbf{p}}_I$ is set to zero accordingly.

A. Gyroscopes' Biases Initialization

The *complete* zero-velocity update described in Section III-D cannot be directly applied to initialize the gyroscope biases. This is due to the fact that an estimate of the orientation ${}^I\bar{q}_G$, required for evaluating $\mathbf{H}_\zeta$ [see (36)], cannot be obtained before estimating the gyroscope biases (Section IV-B.2). Instead, to provide an initial estimate for the gyroscope biases, $\mathbf{b}_g$, we use *partial* zero-velocity updates. In particular, we initially set $\hat{\mathbf{b}}_g$ to an arbitrary value (e.g., zero), while its covariance is set to a large value, reflecting the lack of *a priori* knowledge about the estimates. Then, we keep the IMU static (i.e., $\boldsymbol{\omega} = \mathbf{0}_{3\times 1}$) and use the second block row of (34)-(36) to perform a partial zero-velocity update. This process is equivalent to averaging the (static) gyroscope measurements to compute an initial estimate of the bias.

B. Orientation Initialization

Since the IMU and the laser scanner are rigidly connected and their relative transformation is known (see Section VI), the initial orientation of the IMU can be directly computed from the initial orientation of the laser scanner. We describe two methods to compute the orientation of the laser scanner using line measurements of three planes with linearly-independent normal vectors. The first method, adapted from Chen (1991), requires observation of all three planes from the same viewpoint, while the second method is capable of using laser scan measurements taken from different perspectives by exploiting the motion information from the gyroscopes.

1) Concurrent observation of three planes: When three non-parallel planes are scanned from the same viewpoint (i.e., the same frame of reference), the estimate of the orientation ${}^I\bar{q}_G$ is initialized using the method of Chen (1991). In this case, three quadratic constraints in terms of the unit quaternion ${}^I\bar{q}_G$ are obtained from the laser scans [see (18)], each of them describing the relationship between a line measurement and the corresponding plane:

$$ z_{1,i} = {}^G\boldsymbol{\pi}_i^T \mathbf{C}^T({}^I\bar{q}_G)\, {}^I\boldsymbol{\ell}_i^\perp = 0, \quad i = 1, \ldots, 3. \tag{37} $$

Chen's algorithm algebraically manipulates the rotation matrix to convert this system of equations to an eighth-order univariate polynomial in one of the d.o.f. of the unknown rotation. Eight solutions for this univariate polynomial are obtained, for example, using the Companion matrix (Cox et al. 2004). The remaining 2 d.o.f. of the rotation, ${}^I\bar{q}_G$, are subsequently determined by back-substitution. In general, an external reference is required to identify the true solution from the eight possibilities. In our work, we employ the gravity measurement from the accelerometers and the planes' identities to find the unique solution.

2) Motion-aided orientation initialization: In order to use Chen's method for initializing the orientation, all three line measurements must be expressed with respect to the same frame of reference; hence three non-parallel planes must be concurrently observed by the laser scanner from the same viewpoint. However, satisfying this prerequisite is quite limiting since it requires facing a corner of a room, for example, where three structural planes intersect. In this work, we address this issue by using the gyroscope measurements to *transform* the laser scans, taken from different viewpoints at different time instants, to a *common frame of reference*. We choose as the common frame, the IMU frame when the first laser scan is recorded (i.e., at time t_1), and denote it by $\{I_1\}$. In this way, we can rewrite the inferred measurement constraints (18) at time $t_j,\ j = 2, 3$ as

$$ {}^G\boldsymbol{\pi}_j^T \mathbf{C}^T({}^I\bar{q}_G(t_j)) {}^I\boldsymbol{\ell}_j^\perp(t_j) = {}^G\boldsymbol{\pi}_j^T \mathbf{C}^T({}^I\bar{q}_G(t_1)) {}^{I_1}\boldsymbol{\ell}_j^\perp(t_j) = 0 \tag{38} $$

where ${}^{I_1}\boldsymbol{\ell}_j^\perp(t_j) = \mathbf{C}({}^{I_1}\bar{q}_{I_j}) {}^I\boldsymbol{\ell}_j^\perp(t_j)$ is the line direction corresponding to the plane Π_j, recorded at time t_j, and transformed to the frame $\{I_1\}$. Since the gyroscope biases are already initialized, the quaternions ${}^{I_1}\bar{q}_{I_j}$ can be obtained by integrating the rotational velocity measurements [see (8) and (6)] between time instants t_1 and t_j. Once all the line

directions, ${}^{I}\boldsymbol{\ell}_j^{\perp}(t_j)$, are expressed with respect to $\{I_1\}$, we employ Chen's algorithm, described before, to find the initial orientation, ${}^{I}\bar{q}_G(t_1)$.

The covariance of the initial orientation estimate is obtained by computing the corresponding Jacobians [by linearizing (38)] and using the uncertainty (covariance) in the estimates of ${}^{I}\boldsymbol{\ell}_j^{\perp}$ and ${}^{I_1}\bar{q}_{I_j}$. However, note that the estimates of the relative transformations ${}^{I_1}\bar{q}_{I_2}$ and ${}^{I_1}\bar{q}_{I_3}$ are correlated. To account for these correlations, we employ the *stochastic cloning* technique (Mourikis et al. 2007) to augment the state vector and the covariance matrix of the EKF with ${}^{I_1}\bar{q}_{I_j}$ at time t_j (assuming we have started integrating from time t_1). In this way, we are able to estimate the IMU orientation by integrating the gyroscope measurements, and concurrently compute the correlations between the IMU orientation estimates at the time instants when laser scans are recorded.

C. Accelerometers' Biases Initialization

In this step, similar to the gyroscope bias initialization, we set the estimate for the accelerometer biases, $\mathbf{b}_a$, to an arbitrary value (e.g., zero), and set its covariance to a sufficiently large value, representing our uncertainty about the arbitrary initial estimate. Since the IMU is initially static, we set the velocity estimate, ${}^{G}\mathbf{v}_I$, and its covariance to zero. Then, keeping the IMU static, we utilize the *complete* zero-velocity update described in Section III-D to initialize the accelerometer biases.

V. Observability Analysis

A key task when designing any estimator is to study the observability properties of the underlying system, to determine if the available measurements will provide enough information to estimate the state. In this section, we prove that the presented system for IMU-laser scanner localization is observable when three known planes (i.e., available from the "as-built" or "as-designed" blueprints), whose normal vectors are linearly independent, are concurrently observed by the laser scanner. Under this condition, which is fulfilled in most practical scenarios (e.g., if the scan plane intersects two walls and the floor), we can employ the pose estimation method described in Section IV to estimate $({}^{G}\mathbf{p}_I, {}^{I}\bar{q}_G)$. For the purpose of observability analysis, we introduce two new *inferred measurements* $\mathbf{h}_1^*$ and $\mathbf{h}_2^*$ that replace the laser scan measurements (18), (22):

$$ {}^{I}\bar{q}_G = \mathbf{h}_1^*(\mathbf{x}) = \boldsymbol{\xi}_1({}^{I}\boldsymbol{\ell}_1, {}^{I}\boldsymbol{\ell}_2, {}^{I}\boldsymbol{\ell}_3) \tag{39} $$

$$ {}^{G}\mathbf{p}_I = \mathbf{h}_2^*(\mathbf{x}) = \boldsymbol{\xi}_2({}^{I}\boldsymbol{\ell}_1, {}^{I}\boldsymbol{\ell}_2, {}^{I}\boldsymbol{\ell}_3). \tag{40} $$

The two functions $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ in (39) and (40) do not need to be known explicitly; only their functional relation with the random variables, ${}^{I}\bar{q}_G$ and ${}^{G}\mathbf{p}_I$, is required for the observability analysis. Our approach uses the Lie derivatives (Hermann and Krener 1977) of the above inferred measurements (39) and (40) for the system in (2)-(5), to show that the corresponding observability matrix is full rank. For this purpose, we first rearrange the nonlinear kinematic equations (2)-(5) in a suitable form for computing the Lie derivatives:

$$ \begin{bmatrix} {}^{I}\dot{\bar{q}}_G \\ \dot{\mathbf{b}}_g \\ {}^{G}\dot{\mathbf{v}}_I \\ \dot{\mathbf{b}}_a \\ {}^{G}\dot{\mathbf{p}}_I \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{1}{2}\boldsymbol{\Xi}({}^{I}\bar{q}_G)\mathbf{b}_g \\ \mathbf{0}_{3\times1} \\ {}^{G}\mathbf{g} - \mathbf{C}^T({}^{I}\bar{q}_G)\mathbf{b}_a \\ \mathbf{0}_{3\times1} \\ {}^{G}\mathbf{v}_I \end{bmatrix}}_{\mathbf{f}_0} + \underbrace{\begin{bmatrix} \frac{1}{2}\boldsymbol{\Xi}({}^{I}\bar{q}_G) \\ \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} \end{bmatrix}}_{\underline{\mathbf{f}}_1} \boldsymbol{\omega}_m + \underbrace{\begin{bmatrix} \mathbf{0}_{4\times3} \\ \mathbf{0}_{3\times3} \\ \mathbf{C}^T({}^{I}\bar{q}_G) \\ \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} \end{bmatrix}}_{\underline{\mathbf{f}}_2} \mathbf{a}_m \,, \tag{41} $$

where $\boldsymbol{\omega}_m$ and $\mathbf{a}_m$ are considered the control inputs, and

$$ \boldsymbol{\Xi}(\bar{q}) \triangleq \begin{bmatrix} q_4\mathbf{I}_3 + \lfloor \mathbf{q} \times \rfloor \\ -\mathbf{q}^T \end{bmatrix} \quad \text{with} \quad \bar{q} = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix}. \tag{42} $$

Note also that $\mathbf{f}_0$ is a 16×1 vector, while $\underline{\mathbf{f}}_1$ and $\underline{\mathbf{f}}_2$ are matrices of dimensions 16×3.

In order to prove that the system is locally weakly observable, it *suffices* to show that the observability matrix, whose rows comprise the gradients of the Lie derivatives of the measurements $\mathbf{h}_1^*$ and $\mathbf{h}_2^*$ with respect to $\mathbf{f}_0$, $\underline{\mathbf{f}}_1$, and $\underline{\mathbf{f}}_2$ [see (41)], is full rank (Hermann and Krener 1977). Since the measurement and kinematic equations describing the IMU-laser scanner localization are infinitely smooth, the observability matrix has an infinite number of rows. However, to prove it is full rank, it suffices to show that a subset of its rows are linearly independent. The following matrix contains one such subset of rows whose linear independence can be easily shown using block Gaussian elimination (Mirzaei and Roumeliotis 2009):

$$ \begin{bmatrix} \nabla\mathfrak{L}_{\mathbf{f}_0}^0\mathbf{h}_1^* \\ \nabla\mathfrak{L}_{\mathbf{f}_0}^0\mathbf{h}_2^* \\ \nabla\mathfrak{L}_{\mathbf{f}_0}^1\mathbf{h}_1^* \\ \nabla\mathfrak{L}_{\mathbf{f}_0}^1\mathbf{h}_2^* \\ \nabla\mathfrak{L}_{\mathbf{f}_0}^2\mathbf{h}_2^* \end{bmatrix} = \begin{bmatrix} \mathbf{I}_4 & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} \\ \mathbf{0}_{3\times4} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_3 \\ \mathbf{X}_1 & -\frac{1}{2}\boldsymbol{\Xi}({}^{I}\bar{q}_G) & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} \\ \mathbf{0}_{3\times4} & \mathbf{0}_{3\times3} & \mathbf{I}_3 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{X}_2 & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{C}^T({}^{I}\bar{q}_G) & \mathbf{0}_{3\times3} \end{bmatrix}. $$

In this matrix, $\mathfrak{L}_{\mathbf{f}_0}^i\mathbf{h}_j^*(\mathbf{x})$ denotes the i-th order Lie derivative of $\mathbf{h}_j^*(\mathbf{x})$ with respect to $\mathbf{f}_0$. The matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ have dimensions 4×4 and 3×4, respectively, and do not need to be computed explicitly since they will be eliminated by the block element $(1,1)$ of the matrix, i.e., the identity matrix $\mathbf{I}_4$. Since $\boldsymbol{\Xi}(\bar{q})$ and $\mathbf{C}(\bar{q})$ are always full rank for any unit quaternion $\bar{q}$ (Mirzaei and Roumeliotis 2009), all the rows of the above matrix are linearly independent. Hence, we conclude the observability analysis with the following lemma:

Lemma 1: Given line measurements corresponding to three known planes with linearly independent normal vectors, the system describing the IMU-laser scanner localization is locally weakly observable.

Simply put, as long as the laser scanner measures the walls, as well as the floor or ceiling, the filter should be able to maintain an accurate estimate the pose of the person. As the person moves through the environment, the laser scanner measures different planes over time, leading to higher accuracy estimates. When the sensing platform stops moving, we can apply zero velocity updates (see Section III-D), to reduce drift.

VI. IMU-Laser Scanner Extrinsic Calibration

The laser scan measurements must be transformed to the IMU frame before an EKF update can be performed. In

particular, in the orientation constraint (18), the measured line direction ${}^{L}\boldsymbol{\ell}^{\perp}$ that is registered in the laser scan frame, is expressed with respect to the IMU frame, $\{I\}$. Similarly, in the distance constraint (22), the perpendicular vector to the line direction, ${}^{L}\boldsymbol{\ell}$, is first transformed to the IMU frame. To perform these transformations, we need to know $({}^{I}\bar{q}_L, {}^{I}\mathbf{p}_L)$, i.e., the rotation and translation between the IMU frame and the laser frame. If the transformation between the IMU and the laser scanner is not precisely known, the constraints (18) and (22) will not hold, and updating the filter based on them can result in inconsistency and divergence of the estimator.

Some methods exist in the literature for extrinsic laser scanner calibration (e.g., Skaloud and Lichti (2006); Rieger et al. (2010)), however, these have primarily focused on recovering the relative orientation of the sensor (i.e., roll, pitch, and yaw angles), and utilize GPS as an additional aid in the calibration process. In contrast, we seek to compute the frame transformation between the laser and IMU using only the sensors' own motion and measurements to planes in the environment.

To address this issue, we have employed a method similar to our previous work for IMU-camera calibration (Mirzaei and Roumeliotis 2008) to calibrate the transformation between the IMU and the laser scanner. For this purpose, we have included $({}^{I}\bar{q}_L, {}^{I}\mathbf{p}_L)$ in the state vector of the EKF, i.e.,

$$\begin{aligned}\mathbf{x}^{aug} &= \begin{bmatrix} {}^{I}\bar{q}_G^T & \mathbf{b}_g^T & {}^{G}\mathbf{v}_I^T & \mathbf{b}_a^T & {}^{G}\mathbf{p}_I^T & | & {}^{I}\bar{q}_L^T & {}^{I}\mathbf{p}_L^T & | & d_1 \cdots d_N \end{bmatrix}^T \\ &= \begin{bmatrix} \mathbf{x}_s^T & | & \mathbf{x}_c^T & | & \mathbf{x}_d^T \end{bmatrix}^T. \end{aligned} \quad (43)$$

We augment the system equations (2)-(5) with

$$ {}^{I}\dot{\bar{q}}_L^T = 0 \quad , \quad {}^{I}\dot{\mathbf{p}}_L^T = 0 \quad (44)$$

which specify that the IMU-laser transformation is rigid and does not change with time. We also extend (20) and (24) to include the corresponding Jacobians with respect to the ${}^{I}\mathbf{p}_L$ and ${}^{I}\bar{q}_L$, respectively. We do so by first writing the orientation and distance constraints explicitly in terms of the laser-to-IMU transformation parameters $({}^{I}\bar{q}_L, {}^{I}\mathbf{p}_L)$, i.e.,

$$z_1 = {}^{G}\boldsymbol{\pi}_i^T \, \mathbf{C}^T({}^{I}\bar{q}_G) \, \mathbf{C}({}^{I}\bar{q}_L) \, {}^{L}\boldsymbol{\ell}^{\perp} = 0 \quad (45)$$

$$z_2 = {}^{G}\boldsymbol{\pi}_i^T \left({}^{G}\mathbf{p}_I + \mathbf{C}^T({}^{I}\bar{q}_G) \left({}^{I}\mathbf{p}_L + \rho \, \mathbf{C}({}^{I}\bar{q}_L) {}^{L}\boldsymbol{\ell} \right) \right) - d_i = 0. \quad (46)$$

The linearized error models for (45) and (46) are

$$\begin{aligned}\widetilde{z}_1 &\simeq \mathbf{h}_{1,s}^T \, \widetilde{\mathbf{x}}_s + \mathbf{h}_{1,d}^T \, \widetilde{\mathbf{x}}_d + \boldsymbol{\gamma}_1^T \, \mathbf{n}_\ell \\ &\quad + \begin{bmatrix} {}^{G}\boldsymbol{\pi}_i^T \mathbf{C}^T\left({}^{I}\hat{\bar{q}}_G\right) \lfloor \mathbf{C}\left({}^{I}\hat{\bar{q}}_L\right) {}^{L}\boldsymbol{\ell}_m^{\perp} \times \rfloor & \mathbf{0}_{1\times 3} \end{bmatrix} \widetilde{\mathbf{x}}_c \\ &= \mathbf{h}_{1,s}^T \, \widetilde{\mathbf{x}}_s + \mathbf{h}_{1,c}^T \, \widetilde{\mathbf{x}}_c + \mathbf{h}_{1,d}^T \, \widetilde{\mathbf{x}}_d + \boldsymbol{\gamma}_1^T \, \mathbf{n}_\ell, \end{aligned} \quad (47)$$

$$\begin{aligned}\widetilde{z}_2 &\simeq \mathbf{h}_{2,s}^T \, \widetilde{\mathbf{x}}_s + \mathbf{h}_{2,d}^T \, \widetilde{\mathbf{x}}_d + \boldsymbol{\gamma}_2^T \, \mathbf{n}_\ell \\ &\quad + \begin{bmatrix} {}^{G}\boldsymbol{\pi}_i^T \mathbf{C}^T\left({}^{I}\hat{\bar{q}}_G\right) \lfloor \mathbf{C}\left({}^{I}\hat{\bar{q}}_L\right) \rho \, {}^{L}\boldsymbol{\ell}_m \times \rfloor & {}^{G}\boldsymbol{\pi}_i^T \mathbf{C}^T\left({}^{I}\hat{\bar{q}}_G\right) \end{bmatrix} \widetilde{\mathbf{x}}_c \\ &= \mathbf{h}_{2,s}^T \, \widetilde{\mathbf{x}}_s + \mathbf{h}_{2,c}^T \, \widetilde{\mathbf{x}}_c + \mathbf{h}_{2,d}^T \, \widetilde{\mathbf{x}}_d + \boldsymbol{\gamma}_2^T \, \mathbf{n}_\ell, \end{aligned} \quad (48)$$

where the calibration error-state is $\widetilde{\mathbf{x}}_c = \begin{bmatrix} {}^{I}\boldsymbol{\delta\theta}_L & {}^{I}\widetilde{\mathbf{p}}_L^T \end{bmatrix}^T$, the Jacobians with respect to the state and line parameters, $\mathbf{h}_{i,s}^T$, $\mathbf{h}_{i,d}^T$, $\boldsymbol{\gamma}_i^T$, $i = 1, 2$, are defined in (20) and (24), and the Jacobians with respect to the calibration parameters, $\mathbf{h}_{i,c}^T$, $i = 1, 2$ are implicitly defined in (47) and (48).

The key idea for IMU-laser calibration is to estimate the augmented state $\mathbf{x}^{aug}$ while in a known or unknown environment with at least three perpendicular walls. Note that since there is not enough information to estimate the calibration from a single viewpoint, we must employ a "motion-induced" calibration strategy. In particular, based on a Lie derivative analysis of the system observability properties (see Section V, Mirzaei and Roumeliotis (2008, 2009)), we have shown that the IMU-laser calibration parameters are observable when at least two rotations are performed about different axes, but we omit the details here for brevity. We move the sensor package and collect data until a satisfactory level of accuracy for the calibration parameters (based on the 3σ bounds computed from the estimated covariance matrix) has been achieved. The results of our on-line calibration process, obtained while exploring an unknown area, are presented in Section VII-C.

VII. Experimental Results

Our proposed IMU-laser localization and mapping algorithm was evaluated with a sensing package comprised of a solid-state ISIS IMU operating at 100 Hz and a SICK LMS200 laser scanner operating at 10 Hz, mounted on a navigation box to log data. These sensors were interfaced to a laptop via RS-232 which recorded the time-stamped measurements. The data-logging software was implemented in C++, whereas the EKF was written in MATLAB.

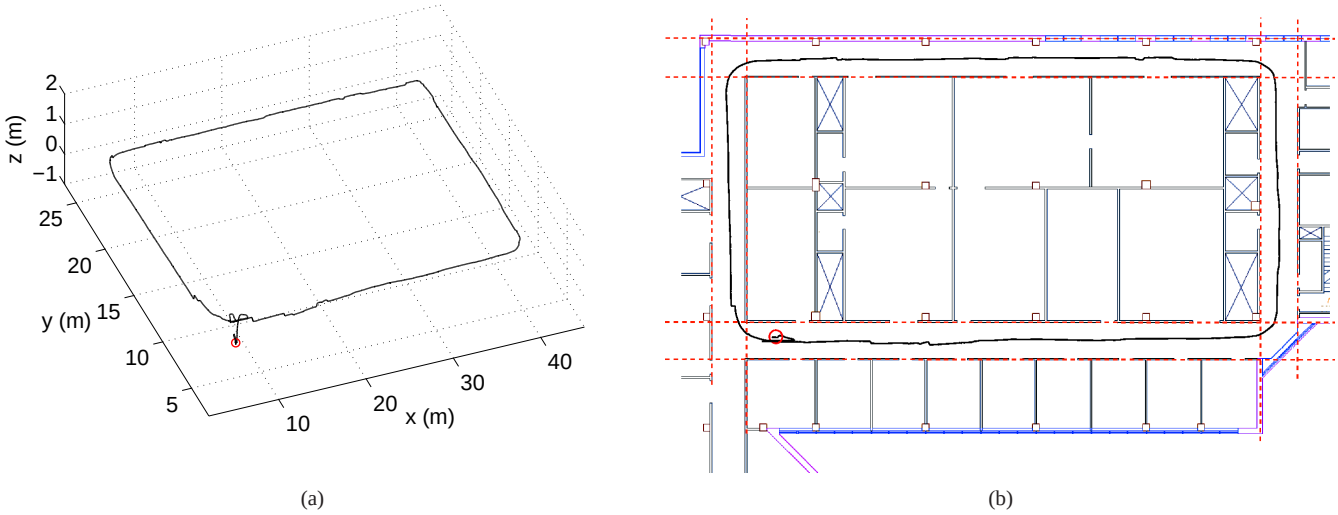
A. Navigation in a known environment

During the first experiment we tested the navigation algorithm in a known environment along a trajectory loop of 120 m in length[3]. The motion profile of the sensor platform contained instantaneous stationary time periods to allow for zero-velocity updates. These updates cause small reductions in the position estimates' covariance [see Fig. 5(a)]. Larger reductions in the covariance take place whenever the laser scanner detects three planes whose normal vectors are linearly independent (e.g., two perpendicular walls and the ceiling) within a short period of time; an event that typically occurs at hallway intersections (e.g., t=49 sec). The *a priori* known map, available from the building blueprints, contained 9 walls and the ceiling. Employing this map, nearly $12,000$ measurement updates were performed during the 8.5 minute trial. The combination of the laser measurements and zero-velocity updates allowed the filter to maintain a precise pose estimate of the sensor platform. Specifically, the maximum uncertainty in the position estimates was 9.16 cm (1σ), while the maximum uncertainty in the attitude estimates was 0.1 deg (1σ) [see Fig. 5(a) and Fig. 5(b)]. The final position uncertainty was $\begin{bmatrix} 27.5 & 1.2 & 1.3 \end{bmatrix}$ cm (3σ). Note that the x-direction uncertainty is larger in the final corridor, since no planes are observed that provide information along the x-axis.

B. Navigation in a previously unknown environment

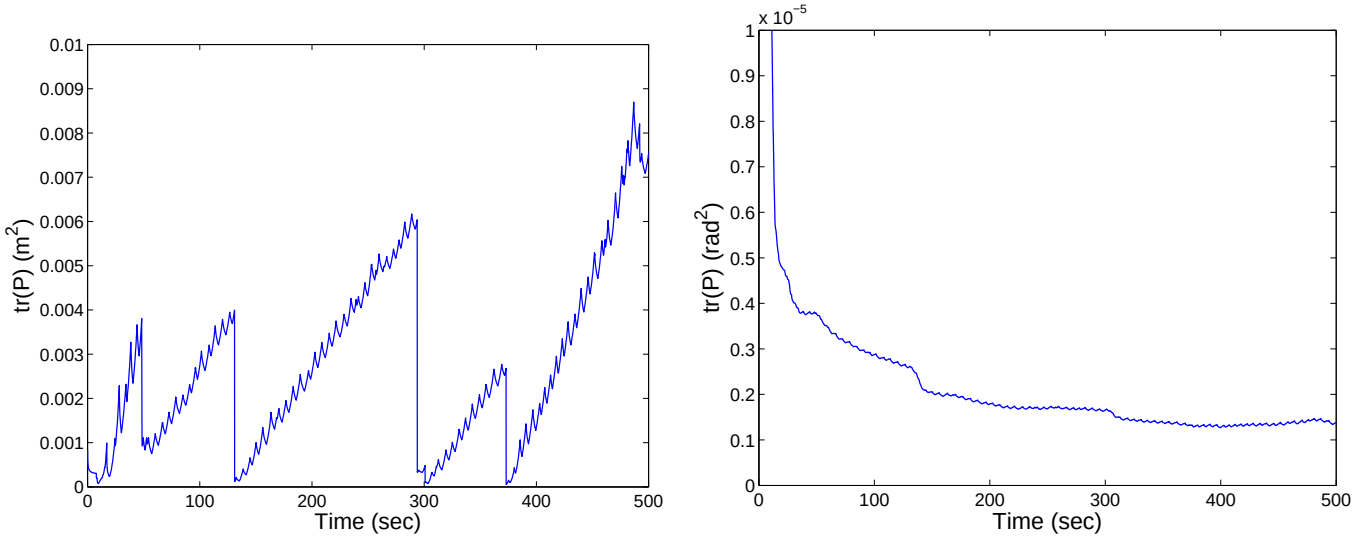
We conducted a second experiment in a previously unknown indoor environment, along a closed-loop path of approximately 270 m in length (see Fig. 4(a) and Fig. 4(b)). The 3D

[3] Video available at `http://mars.cs.umn.edu/videos/IMU-Laser.m4v`



(a) (b)

Fig. 2. (a) 3D view of the estimated trajectory. The sensing package was initially placed on the ground for the purpose of IMU-bias initialization, and subsequently picked up and carried in a clock-wise loop of 120 m in length through the building hallways. (b) Top-view of the estimated 3D trajectory during an 8.5 min experiment. The red circle indicates the starting position (on the floor), and the dashed red lines indicate the walls which were included in the building map.

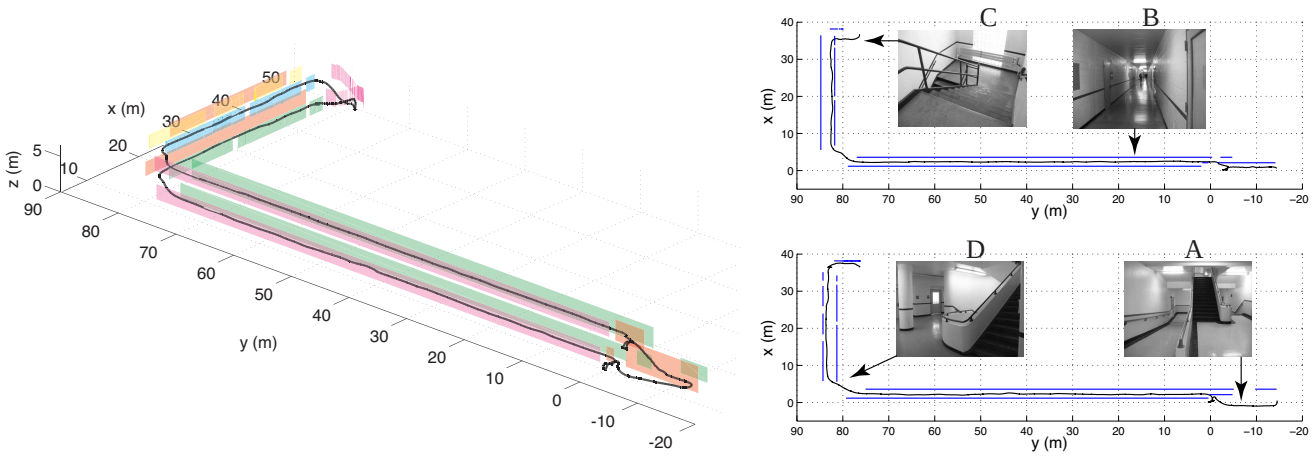


(a) (b)

Fig. 3. (a) The trace of the position covariance. During the run, the maximum uncertainty along any axis was 9.16 cm. (1σ). (b) The trace of the attitude covariance. During the run, the maximum uncertainty about any axis was 0.1 deg. (1σ).

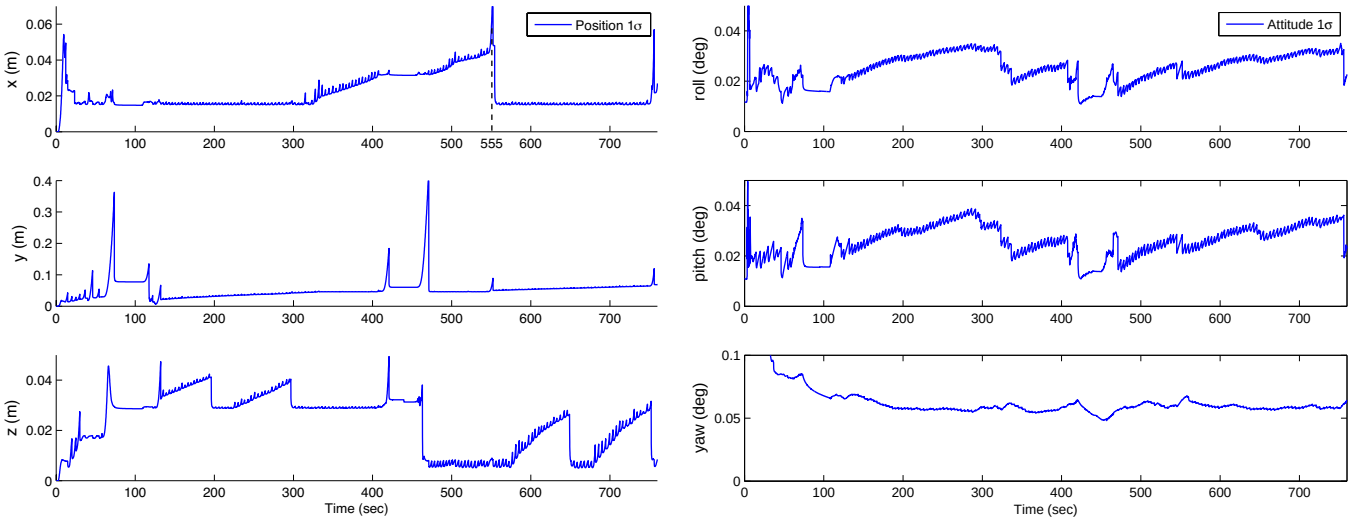
trajectory covered two floors of Akerman Hall at the University of Minnesota, which included traversing two stairways and a ramp. The environment contained a multitude of clutter (e.g., trash cans, open and closed doors, storage boxes, and furniture), as well as normal pedestrian traffic flow. Despite the large amount of non-planar objects observed by the laser scanner, our localization aid accurately captured the 3D layout of the building, which in turn enabled precise localization.

During the experiment, as in the known map case, the motion profile of the sensor platform contained instantaneous stationary time periods to allow for zero-velocity updates. These updates caused small reductions in the position estimates' covariance [see Fig. 5(a)]. Larger reductions in the covariance occurred whenever an estimated structural plane was re-detected (e.g., $t = 555$ sec, x-axis update). The trajectory was accurately tracked, with an average position uncertainty of 3.18 cm (1σ), and an average attitude uncertainty of 0.02 deg (1σ) [see Figs. 5(a), 5(b)]. The final position uncertainty after loop closure was $\begin{bmatrix}2.29 & 6.84 & 0.43\end{bmatrix}$ cm (1σ). In addition to tracking the six d.o.f. pose of the person, a map was constructed which contained 16 walls and the ceilings of both building levels (see Figs. 4(a) and 4(b)). The uncertainty of the least accurately estimated distance to a wall was 4.57 cm (1σ), whereas the average uncertainty for all planes was 1.51 cm (1σ). The quality of the map and trajectory estimates is due to more than $19,000$ measurement updates that were performed



(a) (b)

Fig. 4. (a) As the person walks with the sensing package, the filter estimates their 3D trajectory as well as a 3D representation of the unknown environment comprised of planar features. A side-view of the estimated 270 m trajectory is shown, which covers two floors of the building. The estimated walls on the first and second floors are depicted, but the estimated ceiling and floor planes have been omitted for clarity of presentation. (b) A top-view of the estimated 3D trajectory during the 13 min experiment. The total length of the trajectory is 270 m. The trajectory starts on the first floor (bottom figure), climbs up the disability ramp and the front stairs (picture A), and traverses the corridors (picture B) of the second floor clockwise (top figure). Subsequently, it descends back to the first floor on the second staircase (picture C), and traverses the first floor (bottom figure) counter clockwise, returning to the origin. Picture D shows the *curved* intersection of the two corridors where no wall was detected. The estimated walls are depicted in blue, and the ceiling and floor have been omitted for clarity of presentation.



(a) (b)

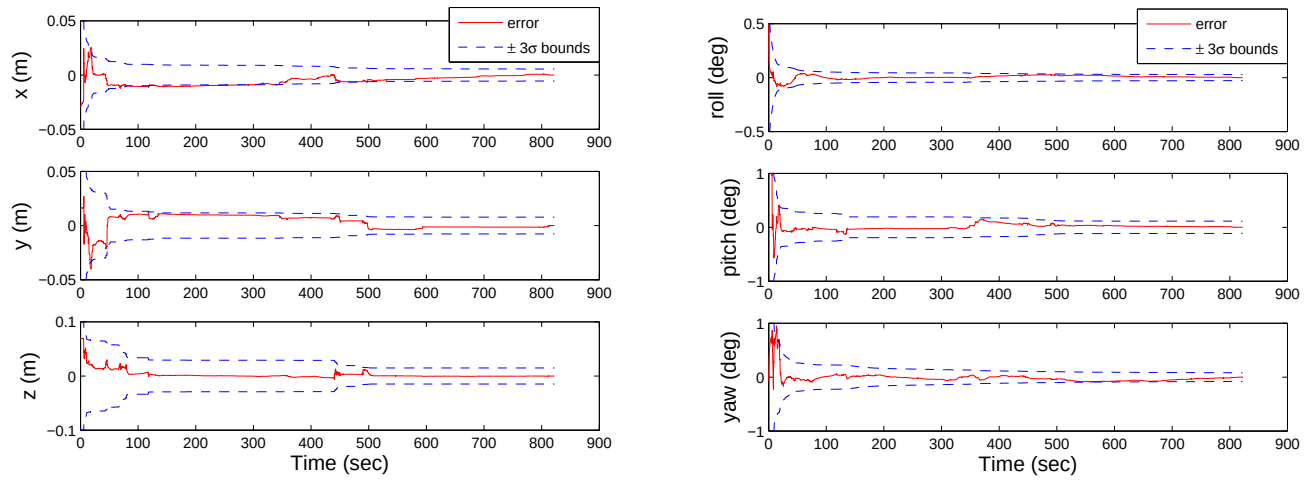
Fig. 5. (a) The 1σ for the x, y, and z axes. During the run, the maximum uncertainty along any axis was 43.94 cm (1σ), while the average 1σ for the least accurate axis was 5.16 cm. (b) The 1σ for the roll, pitch, and yaw angles computed from the angle-error covariance. During the run, the maximum uncertainty about any axis was 0.06 deg. (1σ).

during the 13 minute trial.

C. Extrinsic laser-to-IMU calibration

We now present the results of our extrinsic laser-to-IMU calibration process. Following the procedure of Section VI, we augmented the state vector with the laser-to-IMU transformation $\{{}^{I}\bar{q}_{L}, {}^{I}\mathbf{p}_{L}\}$, and concurrently estimated these parameters while navigating in a previously unknown building (see Section VII-B). We note that calibration can be made more accurate and converge faster if completed during a separate initialization phase in an environment with perfectly known planes; however, our algorithm performs accurately in both scenarios.

Figures 6(a) and 6(b) depict the results for the position and orientation estimates, respectively. In order to demonstrate the consistency of the calibration process, we compute the error of the estimates with respect to the final estimate, along with the corresponding 3σ bounds. We note that since this is an experimental trial, it is impossible to know the true value of the rotation and translation between the laser and IMU; however, the obtained results match closely with the best estimates that we could achieve through hand-measured



(a) (b)

Fig. 6. (a) The relative-translation error (computed versus the final estimate) and the corresponding 3σ bounds for the laser-to-IMU translation vector. The final uncertainties were 0.54 cm along x, 0.76 cm along y, and 1.47 cm along z (3σ). The final translation estimate was ${}^{I}\mathbf{p}_{L} = [25.91 \quad -3.13 \quad -13.42]^{T}$ cm, which agrees with our best hand-measured estimates. (b) The relative-orientation error (computed versus the final estimate) and the corresponding 3σ bounds for the laser-to-IMU rotation ${}^{I}\bar{q}_{L}$. The final uncertainties were 0.02 deg in roll, 0.11 deg in pitch, and 0.08 deg in yaw (3σ). The final orientation estimate was 177.44 deg in roll, 67.4 deg in pitch, and -2.29 deg in yaw (converted from quaternion to roll-pitch-yaw convention), which agrees with our best hand-measured estimates.

techniques. The estimated laser-to-camera translation vector was ${}^{I}\mathbf{p}_{L} = \begin{bmatrix}25.91 & -3.13 & -13.42\end{bmatrix}$ cm, and the estimated orientation was 177.44 deg in roll, 67.4 deg in pitch, and -2.29 deg in yaw, which we converted from quaternion to roll-pitch-yaw convention for ease of presentation. The most uncertain axis for position was 1.47 cm (3σ) along z, while the most uncertain axis for orientation was 0.11 deg (3σ) about y.

VIII. Conclusions and Future Work

This paper presented a novel L-INS, based on a 2D laser scanner and an IMU, capable of 3D localization and mapping in indoor environments. In the proposed method, the orthogonal structural planes of the building are employed as landmarks to aid in localization. Since the building layout may be partially or completely unknown, the planes' parameters are estimated concurrently with the six d.o.f. pose of the person. To this end, an EKF is utilized to fuse information from an IMU and a 2D laser scanner, and estimate the person's motion, and the building's structural planes. We presented a practical method for filter initialization using line-to-plane correspondences to initialize the orientation and zero-velocity updates to initialize the IMU bias estimates. Furthermore, we studied the observability properties of the system to determine a sufficient condition on the number and type of measurements so as to ensure the pose can be estimated. As a final contribution of this paper, we proposed a laser-to-IMU calibration method which is capable of on-line estimation of the laser-to-IMU transformation. The validity of the proposed method is demonstrated in experimental trials in both previously known and unknown environments, which include challenging 3D building structures such as staircases, a disability access ramp, and long corridors. Furthermore, the environments contained a typical amount of office clutter (e.g., chairs and desks) as well as pedestrian traffic.

Our future work includes providing an efficient and intuitive system interface for a visually impaired person. Classification of the non-planar objects and obstacles by processing the laser scans is also within our near goals.

References

J. Borenstein, L. Ojeda, and S. Kwanmuang. Heuristic reduction of gyro drift for personnel tracking systems. *Journal of Navigation*, 62(1):41–58, Jan. 2009.

D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg. Globally consistent 3d mapping with scan matching. *Robotics and Autonomous Systems*, 56(2):130 – 142, 2008.

F. Cavallo, A. M. Sabatini, and V. Genovese. A step toward GPS/INS personal navigation systems: real-time assessment of gait by foot inertial sensing. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1187–1191, Edmonton, Canada, Aug. 2–6, 2005.

H. H. Chen. Pose determination from line-to-plane correspondences: existence condition and closed-form solutions. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(6):530–541, June 1991.

D. M. Cole and P. M. Newman. Using laser range data for 3D SLAM in outdoor environments. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 1556–1563, Orlando, FL, May 15–19, 2006.

D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer, 2004.

M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba. A solution to the simultaneous localization and map building (SLAM) problem. *IEEE Trans. on Robotics and Automation*, 17(3):229–241, June 2001.

S. Ertan, C. Lee, A. Willets, H. Tan, and A. Pentland. A

wearable haptic navigation guidance system. In *Proc. of the Int. Sym. on Wearable Computers*, pages 164–165, Pittsburgh, PA, Oct. 19–20, 1998.

D. Hähnel, W. Burgard, and S. Thrun. Learning compact 3D models of indoor and outdoor environments with a mobile robot. *Robotics and Autonomous Systems*, 44(1):15–27, 2003.

R. Hermann and A. Krener. Nonlinear controlability and observability. *IEEE Trans. on Automatic Control*, 22(5): 728–740, Oct. 1977.

J. A. Hesch and S. I. Roumeliotis. Design and analysis of a portable indoor localization aid for the visually impaired. *Int. Journal of Robotics Research*, 29(11):1400–1415, Sept. 2010.

J. A. Hesch, F. M. Mirzaei, G. L. Mariottini, and S. I. Roumeliotis. A 3D pose estimator for the visually impaired. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 2716–2723, St. Louis, MO, Oct. 11–15, 2009.

J. A. Hesch, F. M. Mirzaei, G. L. Mariottini, and S. I. Roumeliotis. A Laser-aided Inertial Navigation System (L-INS) for human localization in unknown indoor environments. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 5376–5382, Anchorage, AK, May 3–8, 2010.

A. Hub, J. Diepstraten, and T. Ertl. Design and development of an indoor navigation and object identification system for the blind. In *Proc. of the Int. ACM SIGACCESS Conf. on Computers and Accessibility*, pages 147–152, Atlanta, GA, Oct. 18–20, 2004.

Humanware. Trekker talking GPS, 2010. URL `www.humanware.com.[Accessed:Feb.2,2010]`.

L. Iocchi and S. Pellegrini. Building 3D maps with semantic elements integrating 2D laser, stereo vision and IMU on a mobile robot. In *Proc. of the 2nd ISPRS Int. Workshop 3D-ARCH*, Zurich, Switzerland, July 12–13, 2007.

J. Kim and S. Sukkarieh. Real-time implementation of airborne inertial-SLAM. *Robotics and Autonomous Systems*, 55(1):62–71, 2007.

P. Kohlhepp, P. Pozzo, M. Walther, and R. Dillman. Sequential 3D-SLAM for mobile action planning. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 722–729, Sendai, Japan, Sept. 28–Oct. 2, 2004.

V. Kulyukin, C. Gharpure, J. Nicholson, and S. Pavithran. RFID in robot-assisted indoor navigation for the visually impaired. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1979–1984, Sendai, Japan, Sept. 28–Oct. 2, 2004.

E. J. Lefferts, F. L. Markley, and M. D. Shuster. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5(5):417–429, Sept.–Oct. 1982.

D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. Journal of Computer Vision*, 60(2):91–110, Nov. 2004.

H. Makino, I. Ishii, and M. Nakashizuka. Development of navigation system for the blind using GPS and mobile phone combination. In *Proc. of the IEEE Int. Conf. of the Engineering in Medicine and Biology Society*, pages 506–507, Amsterdam, Netherlands, Oct. 31–Nov. 3, 1996.

F. M. Mirzaei and S. I. Roumeliotis. A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE Trans. on Robotics*, 24(5):1143–1156, Oct. 2008.

F. M. Mirzaei and S. I. Roumeliotis. IMU-laser scanner localization: Observability analysis. Technical report, Dept. of Computer Science & Engineering, University of Minnesota, MARS Lab, Minneapolis, MN, Jan. 2009. URL `http://mars.cs.umn.edu/tr/reports/Mirzaei09.pdf`.

A. I. Mourikis and S. I. Roumeliotis. A dual-layer estimator architecture for long-term localization. In *Proc. of the Workshop on Visual Localization for Mobile Platforms*, pages 1–8, Anchorage, AK, June 2008.

A. I. Mourikis, S. I. Roumeliotis, and J. W. Burdick. SC-KF mobile robot localization: A stochastic cloning Kalman filter for processing relative-state measurements. *IEEE Trans. on Robotics*, 23(4):717–730, Aug. 2007.

V. Nguyen, A. Harati, A. Martinelli, R. Siegwart, and N. Tomatis. Orthogonal SLAM: a step toward lightweight indoor autonomous navigation. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 5007–5012, Beijing, China, Oct. 9–15, 2006.

V. Nguyen, S. Gächter, A. Martinelli, N. Tomatis, and R. Siegwart. A Comparison of Line Extraction Algorithms using 2D Range Data for Indoor Mobile Robotics. *Autonomous Robots*, 23(2):97–111, Aug. 2007.

A. Nüchter, H. Surmann, K. Lingemann, J. Hertzberg, and S. Thrun. 6D SLAM with an application in autonomous mine mapping. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 1998–2003, New Orleans, LA, Apr. 18–May 1 2004.

S. T. Pfister, S. I. Roumeliotis, and J. W. Burdick. Weighted line fitting algorithms for mobile robot map building and efficient data representation. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 1304–1311, Taipei, Taiwan, Sept. 14–19, 2003.

L. Ran, S. Helal, and S. Moore. Drishti: an integrated indoor/outdoor blind navigation system and service. In *Proc. of the IEEE Conf. on Pervasive Computing and Communications*, pages 23–30, Orlando, FL, Mar. 14–17, 2004.

P. Rieger, N. Studnicka, M. Pfennigbauer, and G. Zach. Boresight alignment method for mobile laser scanning systems. *Journal of Applied Geodesy*, 4(1):13–21, 2010.

K. Sagawa, H. Inooka, and Y. Satoh. Non-restricted measurement of walking distance. In *Proc. of the IEEE Int. Conf. on Systems, Man, and Cybernetics*, pages 1847–1852, Nashville, TN, Oct. 8–11, 2000.

Sendero. Braille sense nav, 2010. URL `www.senderogroup.com.[Accessed:Feb.2,2010]`.

J. Skaloud and D. Lichti. Rigorous approach to bore-sight self-calibration in airborne laser scanning. *ISPRS Journal of Photogrammetry & Remote Sensing*, 61(1):47–59, Oct. 2006.

R. Smith, M. Self, and P. Cheeseman. Estimating uncertain spatial relationships in robotics. In *Autonomous Robot*

Vehicles, pages 167–193. Springer-Verlag, New York, NY, 1990.

S. Thrun, D. Fox, and W. Burgard. Monte Carlo localization with mixture proposal distribution. In *Proc. of the AAAI National Conference on Artificial Intelligence*, pages 859–865, Austin, TX, July 30–Aug. 3, 2000.

B. Tjan, P. Beckmann, N. Giudice, and G. Legge. Digital sign system for indoor wayfinding for the visually impaired. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition - Workshop on Computer Vision Applications for the Visually Impaired*, San Diego, CA, June 20–25, 2005.

N. Trawny and S. I. Roumeliotis. Indirect Kalman filter for 3D pose estimation. Technical report, MARS Lab, Dept. of Computer Science & Engineering, University of Minnesota, Minneapolis, MN, Mar. 2005.

I. Ulrich and J. Borenstein. The GuideCane - applying mobile robot technologies to assist the visually impaired. *IEEE Trans. on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 31(2):131–136, Mar. 2001.

D. Yuan and R. Manduchi. Dynamic environment exploration using a virtual white cane. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 243–249, San Diego, CA, June 20–25, 2005.