A Laser-Aided Inertial Navigation System (L-INS) for Human Localization in Unknown Indoor Environments

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

Dept. of Computer Science and Engineering, University of Minnesota

Email: {joel|faraz|gianluca|stergios}@cs.umn.edu

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 person and a 3D map of the environment. The IMU measurements are integrated to obtain pose estimates, which are subsequently corrected using line-to-plane correspondences between linear segments in the laser-scan data and orthogonal structural planes of the building. Exploiting the orthogonal building planes ensures fast and efficient initialization and estimation of the map features while providing human-interpretable layout of the environment. The L-INS is experimentally validated by a person traversing a multistory building, and the results demonstrate the reliability and accuracy of the proposed method 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 Trekker) rely on GPS signals and cannot be utilized indoors, under tree cover, or next to tall buildings where reception is poor. In the academic community, numerous indoor and outdoor electronic navigation systems have been proposed. However, the majority of the existing algorithms are designed for mobile robots that are limited to move on planar surfaces [1], [2] or require heavy sensors, such as a 3D laser scanner [3], [4], that cannot be carried by a human. Other algorithms have relied on visual information [5], [6] and as such, they are not robust under variable lighting conditions and
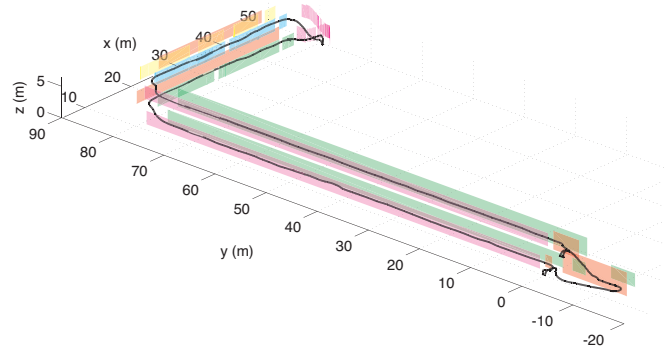
Fig. 1. As the person walks with the sensing package, the filter estimates their 3D trajectory as well as a 3D representation of the environment comprised of planar features. A side-view of the estimated 270 m trajectory is shown, which covers two floors of the building. The estimated walls on the first and second floors are depicted, but the estimated ceiling and floor planes have been omitted for clarity of presentation.

require intensive processing resources which 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 person, allowing them to safely navigate in *3D space* (see Fig. 1).
- The navigation aid should enable the visually impaired person to walk through previously *unknown buildings* without getting lost. This requires the system to build a map of the previously unexplored area and localize with respect to it in *real-time*.
- The sensor package should be *compact, unobtrusive* to the person, and *lightweight* enough to be carried without fatigue.
- The selected sensing modalities 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.

To meet these objectives, we build upon our previous work [7], [8], and present an indoor Laser-aided Inertial Navigation System (L-INS) using an *Inertial Measurement Unit (IMU)* and a *2D laser scanner*. 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, the IMU measurement noise and bias drift would cause the pose estimation errors to grow unbounded over time. To mitigate this issue, we utilize the straight-line features extracted from the 2D laser scans to update the pose estimates. In particular, as the person moves, the laser scanner's attitude changes which allows its scanning plane to intersect a variety of structural planes of the building (i.e., the walls, floor, and ceiling). If the structural planes were known *a priori* as in [7], [8], then the straight-line features could be directly matched with the known map. Unfortunately, in many cases in practice, a building map is not available in advance. To overcome this challenge, we perform *Simultaneous Localization and Mapping (SLAM)* to estimate the layout of the building while concurrently tracking the person's pose. In order to achieve this goal efficiently, we exploit the fact that most indoor structural planes are *orthogonal to each other*. Thus, we fix each plane's orientation the first time it is observed, and only estimate its distance to the origin of the reference frame.

Constructing the map based on orthogonal planar structures has the advantage of keeping the person's orientation error bounded [9] 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 computational complexity 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 a real-world experiment covering multiple levels of an indoor environment. The 270 m trajectory traverses several staircases and a disability access ramp. In addition, the test environment includes significant clutter (e.g., trashcans, storage boxes, and furniture), as well as a normal flow of pedestrian traffic. Despite these challenges, our algorithm accurately tracks the person's path, and correctly estimates a map of the building layout.

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 EKF-based pose estimator. The proposed method is validated with experimental results in Section IV, and we conclude the paper and present future research directions in Section V.

II. RELATED WORK

Most relevant efforts have primarily addressed GPS-based *outdoor navigation* which cannot be used inside a building [10], [11]. *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 [12]. As a person walks, their position is computed by double integration of the acceleration measurements. Unfortunately, the integration of accelerometer bias and noise causes the *position error* to grow unbounded. Even if the rate of position-error increase can be reduced with static-period drift corrections [13], [14], dead-reckoning systems still remain unreliable over long time intervals.

2) Navigating with a known map: Unlike dead-reckoning approaches that do not employ external references, map-based systems infer position and orientation information from known landmarks in the environment. In our previous work [8], we presented a map-based indoor localization aid for the visually impaired using an IMU and a 2D laser scanner. Despite the novelty of our previous approach, it suffers from the same limitations of all map-based localization methods which include: (i) *time and cost* associated with acquiring the map, (ii) the system's *inability to adapt* to spatial layout changes, and (iii) the *restriction* of use to previously mapped areas.

3) Navigating without a pre-existing map: The most flexible navigation aids are those that can exploit environment sensing to perform SLAM. The majority of the proposed systems for SLAM consider either 2D map and sensor motion [15], [16], or restrict the sensor motion to planar surfaces and create a 3D map of the surroundings [1], [2], [9]. However, these algorithms are not 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)] [3], [4], [17], [18], [19]. 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 [5], [6]. The main drawback of camera-based systems is sensitivity to variable lighting conditions, which restricts their use as navigation aids for the visually impaired where reliability is paramount. Additionally, processing images and extracting visual features are typically computationally intensive tasks that are impractical to carry out on hand-held computing devices. Furthermore, the visual landmarks

(e.g., SIFT features [20]) often used in these approaches may not be geometrically meaningful or interpretable for 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. Our system tracks the six d.o.f. pose of the person and maps the 3D structure of the environment, while avoiding the high computational burden associated with processing 3D laser scan data. This sensor pair is ideal since new lightweight models that are robust in various environmental conditions are now commercially available. Furthermore, we use structural planes of the buildings as the map features, which exist in almost all buildings, ensuring 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 the SLAM algorithm do not grow unbounded over time. Finally, we use an EKF to fuse the IMU and laser scanner measurements, which along with the negligible overhead of line-segment extraction from laser data, provides real-time execution even on hand-held computing devices with limited computational and memory resources.

III. Algorithm Description

While walking through an indoor environment, the visually-impaired person is equipped with a navigation aid consisting of an IMU and a 2D laser scanner, which are rigidly connected (see Fig. 2). A hand-held computer collects the measurements from the two sensors, which then are fused in an EKF to concurrently estimate the six d.o.f. pose of the moving platform, as well as the 3D map of the buildings' perpendicular structural planes (i.e., the walls, floor, and ceiling). In what follows, we present the propagation and update models used by the EKF.

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:

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

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. 2), 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 [8]. 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.
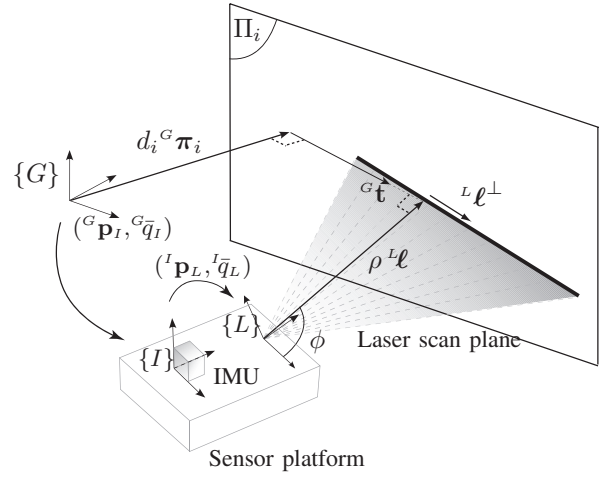

Fig. 2. As the IMU-laser sensor platform moves, the laser scan plane intersects a structural planar surface, Π_i, described by d_i and ${}^G\boldsymbol{\pi}_i$, which are the Hessian normal form components of the plane with respect to the global frame of reference, $\{G\}$. The shortest vector in the 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 known IMU-laser transformation is denoted by $({}^I\mathbf{p}_L, {}^I\bar{q}_L)$, while the IMU pose with respect to $\{G\}$ is $({}^G\mathbf{p}_I, {}^G\bar{q}_I)$.

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, $i = 1, \ldots, N$, 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 $\begin{bmatrix} {}^G\boldsymbol{\pi}_1^T \ldots {}^G\boldsymbol{\pi}_N^T \end{bmatrix}^T$, where each component ${}^G\boldsymbol{\pi}_i$ is determined once during the new plane initialization step (Section III-C). 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 [21], [22]):

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

---

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

In these expressions, $\boldsymbol{\omega}(t) = [\omega_1(t) \quad \omega_2(t) \quad \omega_3(t)]^T$ 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

$$\boldsymbol{\omega}_m(t) = \boldsymbol{\omega}(t) + \mathbf{b}_g(t) + \mathbf{n}_g(t) \tag{6}$$

$$\mathbf{a}_m(t) = \mathbf{C}(^I\bar{q}_G(t)) \left(^G\mathbf{a}(t) - ^G\mathbf{g} \right) + \mathbf{b}_a(t) + \mathbf{n}_a(t), \tag{7}$$

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

$$^I\dot{\hat{\bar{q}}}_G(t) = \frac{1}{2}\boldsymbol{\Omega}(\hat{\boldsymbol{\omega}}(t))^I\hat{\bar{q}}_G(t) \tag{8}$$

$$^G\dot{\hat{\mathbf{p}}}_I(t) = ^G\hat{\mathbf{v}}_I(t) \;,\quad ^G\dot{\hat{\mathbf{v}}}_I(t) = \mathbf{C}^T(^I\hat{\bar{q}}_G(t))\,\hat{\mathbf{a}}(t) + ^G\mathbf{g} \tag{9}$$

$$\dot{\hat{\mathbf{b}}}_g(t) = \mathbf{0}_{3\times1} \;,\quad \dot{\hat{\mathbf{b}}}_a(t) = \mathbf{0}_{3\times1} \tag{10}$$

$$\dot{\hat{d}}_i(t) = 0 \;,\quad i = 1, \ldots, N, \tag{11}$$

with $\hat{\mathbf{a}}(t) = \mathbf{a}_m(t) - \hat{\mathbf{b}}_a(t)$, and $\hat{\boldsymbol{\omega}}(t) = \boldsymbol{\omega}_m(t) - \hat{\mathbf{b}}_g(t)$. The $(15+N) \times 1$ error-state vector is defined as

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

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, $\delta\boldsymbol{\theta}$, implicitly defined by the *error quaternion*

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

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 $E\{\delta\boldsymbol{\theta}\delta\boldsymbol{\theta}^T\}$. Since the attitude corresponds to three d.o.f., this is a minimal representation.

The linearized continuous-time error-state equation is

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

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.,

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

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

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 [22].

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 covariance propagation equation, we evaluate the discrete-time state transition matrix

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

and the discrete-time system noise covariance matrix

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

The propagated covariance is then computed as

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

B. Filter 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. 2). 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.[3] 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

---

[2]We utilized the Split-and-Merge algorithm [23] to segment the laser-scan data and a weighted line-fitting algorithm [24] to estimate the line parameters (ρ, ϕ) for each line.

[3]The *laser-to-IMU rigid transformation* is computed off-line using a laser-to-IMU calibration procedure adapted from [25] to estimate $(^I\mathbf{p}_L, {}^I\bar{q}_L)$.

the plane Π_i, vector ${}^G\boldsymbol{\pi}_i$, is perpendicular to $\mathbf{C}^T({}^I\bar{q}_G)\,{}^I\boldsymbol{\ell}^\perp$ (see Fig. 2), which yields the following *orientation measurement constraint*

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

The expected measurement is

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

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

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

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 [24].

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

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

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. The vector ${}^G\mathbf{t}$ is eliminated by projecting (21) onto ${}^G\boldsymbol{\pi}_i^T$, yielding the *distance measurement constraint*

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

The expected measurement is

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

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

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

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

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

which are used in the expression for the Kalman gain

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

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

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

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

C. New Plane Initialization

In Section III-B, we described the filter update step using a line measurement to one of the N structural planes comprising the estimated map. We now present a procedure to initialize the state and covariance of a previously unknown plane the first time it is observed.[4] 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 considered in the map, $\mathbf{e}_j$, $j = 1, 2, 3$. We employ a Mahalanobis distance test to measure the probability of correspondence, i.e., we compute the orientation residual $r_{1,j} = -\mathbf{e}_j^T\mathbf{C}^T({}^I\hat{\bar{q}}_G)\,{}^I\boldsymbol{\ell}_m^\perp$, $j = 1, 2, 3$, and the covariance of the residual

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

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

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

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.,

$$d' \triangleq \hat{d}_{N+1} = {}^G\boldsymbol{\pi}_{N+1}^T\left({}^G\hat{\mathbf{p}}_I + \mathbf{C}^T({}^I\hat{\bar{q}}_G)\left({}^I\mathbf{p}_L + \rho_m\,{}^I\boldsymbol{\ell}_m\right)\right) \tag{29}$$

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

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

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, $\mathbf{P}_{d'\mathbf{x}}$, i.e.,

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

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

---

[4]Previously detected planes (already in the state vector) are identified using a chi-square test [26].
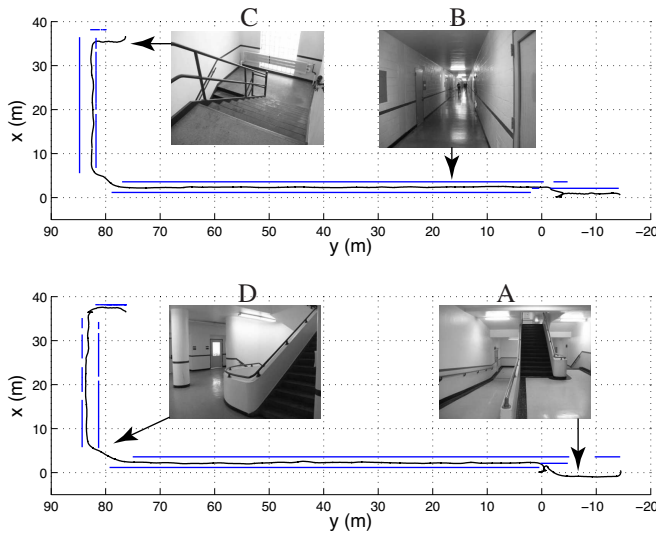
Fig. 3. A top-view of the estimated 3D trajectory during the 13 min experiment. The total length of the trajectory is 270 m. The trajectory starts on the first floor (bottom figure), climbs up the disability ramp and the front stairs (picture A), and traverses the corridors (picture B) of the second 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.

Lastly, the augmented covariance, $\mathbf{P}^{aug}$ is computed as

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

D. Zero-Velocity Update

When the laser scanner does not detect any structural planes along certain directions for an extended period of time, the position estimates accumulate errors along those directions, due to *accelerometer drift*. This effect can be compensated by means of drift correction during instantaneous stationary periods of the motion [13]. The details of this procedure, termed *zero-velocity update*, are given in [8].

IV. 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.[5] 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.

We conducted the experiment in an indoor environment, along a closed-loop path of approximately 270 m in length (see Figs. 1 and 3). The 3D trajectory covered two floors of Akerman Hall at the University of Minnesota, which included traversing two stairways and a ramp. The environment

---

[5]This sensor pair was used for ease of implementation. We are currently replacing these sensors with the small-scale Hokuyo URG laser scanner, and the Memsense nIMU, which can easily fit on a white cane.

contained a multitude of clutter (e.g., trashcans, open and closed doors, storage boxes, and furniture), as well as normal pedestrian traffic flow. Despite the large amount of irrelevant 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, the motion profile of the sensor platform contained instantaneous stationary time periods to allow for zero-velocity updates [8]. These updates caused small reductions in the position estimates' covariance [see Fig. 4(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. 4(a), 4(b)]. The final position uncertainty after loop closure was $\begin{bmatrix} 2.29 & 6.84 & 0.43 \end{bmatrix}$ cm (1σ). This level of accuracy ensures that the system is safe for human navigation indoors, and can be used to help a person avoid hazards such as open stair wells. 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. 1 and 3). 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 during the 13 minute trial. These measurement updates enabled the filter to remain consistent, an observation corroborated by the measurement residuals which lie within their 3σ bounds [see Fig. 4(c)].

V. CONCLUSIONS AND FUTURE WORK

This paper presented a novel L-INS 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 is not known *a priori* 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. The validity of the approach is demonstrated experimentally in a multistory building, on a path that includes staircases, a disability access ramp, and corridors, with a normal flow of 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

[1] S. Thrun, D. Fox, and W. Burgard, "Monte Carlo localization with mixture proposal distribution," in *Proc. of the AAAI National Conference on Artificial Intelligence*, Austin, TX, July 30–Aug. 3, 2000, pp. 859–865.
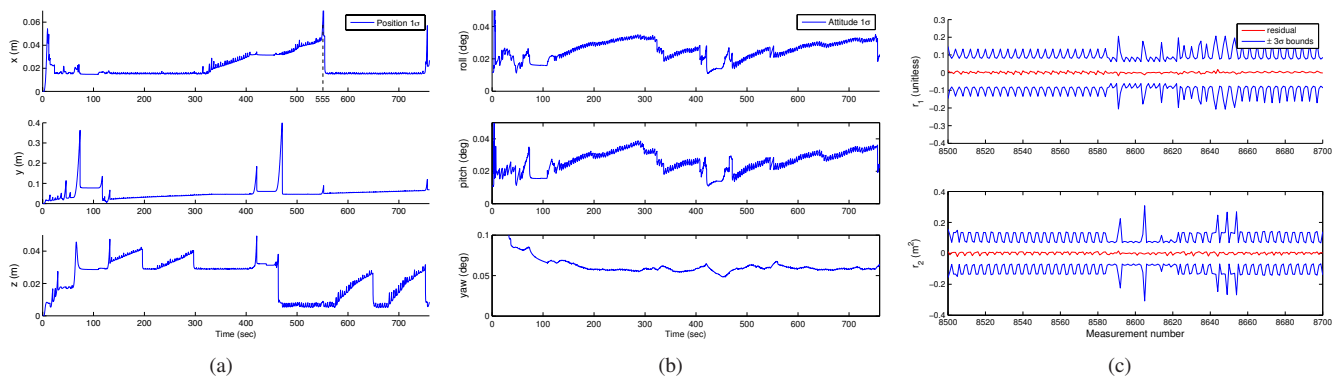
Fig. 4. (a) The 1σ for the x, y, and z axes. During the run, the maximum uncertainty along any axis was 43.94 cm (1σ), while the average 1σ for the least accurate axis was 5.16 cm. (b) The 1σ for the roll, pitch, and yaw angles computed from the angle-error covariance. During the run, the maximum uncertainty about any axis was 0.06 deg. (1σ). (c) More than $19,000$ line measurements were recorded during this experiment. The measurement residuals for a subset of these (200) are plotted in order to demonstrate the filter's consistency.

[2] 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.

[3] D. Hähnel, W. Burgard, and S. Thrun, "Learning compact 3D models of indoor and outdoor environments with a mobile robot," *Robotics and Autonomous Systems*, vol. 44, no. 1, pp. 15–27, 2003.

[4] D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg, "Globally consistent 3d mapping with scan matching," *Robotics and Autonomous Systems*, vol. 56, no. 2, pp. 130 – 142, 2008.

[5] 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*, Anchorage, AK, June 2008, pp. 1–8.

[6] J. Kim and S. Sukkarieh, "Real-time implementation of airborne inertial-SLAM," *Robotics and Autonomous Systems*, vol. 55, no. 1, pp. 62–71, 2007.

[7] J. A. Hesch and S. I. Roumeliotis, "An indoor localization aid for the visually impaired," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, Rome, Italy, Apr. 10–14, 2007, pp. 3545–3551.

[8] 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*, St. Louis, MO, Oct. 11–15, 2009, pp. 2716–2723.

[9] 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*, Beijing, China, Oct. 9–15, 2006, pp. 5007–5012.

[10] 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*, Amsterdam, Netherlands, Oct. 31–Nov. 3, 1996, pp. 506–507.

[11] 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*, Orlando, FL, Mar. 14–17, 2004, pp. 23–30.

[12] 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*, Edmonton, Canada, Aug. 2–6, 2005, pp. 1187–1191.

[13] 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*, Nashville, TN, Oct. 8–11, 2000, pp. 1847–1852.

[14] J. Borenstein, L. Ojeda, and S. Kwanmuang, "Heuristic reduction of gyro drift for personnel tracking systems," *Journal of Navigation*, vol. 62, no. 1, pp. 41–58, Jan. 2009.

[15] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*. New York, NY: Springer-Verlag, 1990, pp. 167–193.

[16] 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*, vol. 17, no. 3, pp. 229–241, June 2001.

[17] 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*, Sendai, Japan, Sept. 28–Oct. 2, 2004, pp. 722–729.

[18] 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*, New Orleans, LA, Apr. 18–May 1 2004, pp. 1998–2003.

[19] 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*, Orlando, FL, May 15–19, 2006, pp. 1556–1563.

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[21] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, Sept.–Oct. 1982.

[22] N. Trawny and S. I. Roumeliotis, "Indirect Kalman filter for 3D pose estimation," MARS Lab, Dept. of Computer Science & Engineering, University of Minnesota, Minneapolis, MN, Tech. Rep., Mar. 2005.

[23] 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*, vol. 23, no. 2, pp. 97–111, Aug. 2007.

[24] 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*, Taipei, Taiwan, Sept. 14–19, 2003, pp. 1304–1311.

[25] 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*, vol. 24, no. 5, pp. 1143–1156, Oct. 2008.

[26] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer-Verlag, 2006.