Design and Analysis of a Portable Indoor Localization Aid for the Visually Impaired

Joel A. Hesch and Stergios I. Roumeliotis

***Abstract*—In this paper, we present the design and analysis of a portable position and orientation (pose) estimation aid for the visually impaired. Our prototype navigation aid consists of a foot-mounted pedometer and a white cane-mounted sensing package, which is comprised of a 3-axis gyroscope and a 2D laser scanner. We introduce a two-layered estimator that tracks the 3D orientation of the white cane in the first layer, and the 2D position of the person holding the cane in the second layer. Our algorithm employs a known building map to update the person's position, and exploits perpendicularity in the building layout as a 3D structural compass. We analytically study the observability properties of the linearized dynamical system, and we provide sufficient observability conditions. We evaluate the real-world performance of our localization aid, and demonstrate its reliability for accurate, real-time human localization.**

***Index Terms*—Handicapped aids, pose estimation, sensor fusion, laser scanner, gyroscopes.**

I. Introduction

MOBILITY is an essential capability for any person who leads an independent life-style. It requires successful execution of several tasks including path planning, navigation, and obstacle avoidance; all of which necessitate accurate assessment of the surrounding environment. For a visually-impaired person such tasks are exceedingly difficult to accomplish, and there are high risks associated with failure in any of these. For this reason, the visually impaired utilize portable navigation aids in their day-to-day activities in order to increase their safety and independence. The standard white cane is one of the most commonly used navigation tools. It is light-weight, inexpensive, and provides tactile cues about potential obstacles, both indoors and outdoors. Unfortunately, white canes cannot provide any information for wayfinding, such as distance to destination, or current heading direction. Assistance dogs are also commonly employed to increase safety when navigating. They can detect and avoid potential hazards, and help someone find their way from one location to another. However, assistance dogs have costly training requirements, and are not typically available to everyone. With the advent of hand-held electronic devices, GPS-based navigation aids have also become commercially available (Sendero Group 2009; Humanware 2009). These are relatively inexpensive and lightweight, but they can only be utilized outdoors in areas where the satellite constellation is visible.

This work was supported by the University of Minnesota (DTC), and the National Science Foundation (IIS-0643680, IIS-0811946, IIS-0835637). Joel Hesch was supported by the NIH Neuro-physical-computational Sciences (NPCS) Graduate Training Fellowship. A preliminary version of this paper was presented at the 2007 IEEE International Conference on Robotics and Automation (Hesch and Roumeliotis 2007).

The authors are with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455, USA (email: {joel|stergios}@cs.umn.edu).

Fig. 1. This figure depicts the typical usage of our localization aid, which estimates the pose of a person walking inside a building. The sensing package is mounted near the handle of the white cane, and a portion of the 2D laser scan plane (60 out of the 240 degrees field of view) is illustrated, along with the line of intersection between the scan plane and the floor.

Unfortunately, portable indoor navigation aids have not been developed with the same success as their outdoor counterparts. The primary barrier is the lack of robust algorithms for determining a person's position and orientation (pose) indoors. An indoor localization aid must rely solely on local features and the building structure in order to infer pose, since an absolute reference system, such as GPS, is not available indoors. This paper presents a novel indoor localization device and algorithms for determining a person's pose inside a building. With the use of this localization aid, guidance and navigation systems can be implemented which will greatly increase the safety and overall mobility of the target demographic.

When designing a portable sensor package for human localization suitable sensor placement must be carefully considered. Body-mounted sensor packages have been presented which require the user to wear an electronic vest or belt fitted with sensing devises (Shoval et al. 1998; Ran et al. 2004). Although mounting a sensor directly on the body simplifies the interpretation of the sensor data (i.e., the transformation from body to sensor is constant and known), it introduces complications

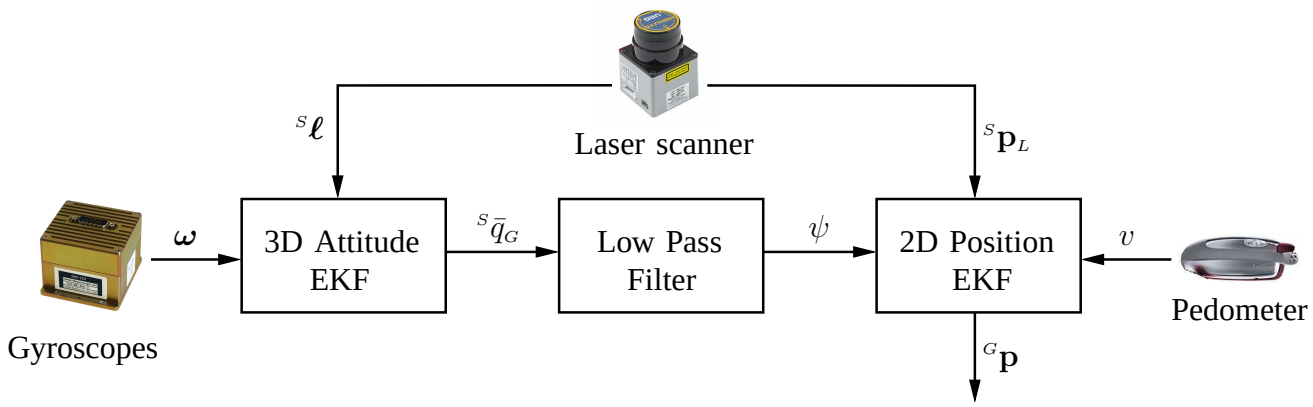

Fig. 2. Algorithm block diagram: The first layer consists of the 3D attitude EKF which utilizes the instantaneous rotational velocity of the cane, $\boldsymbol{\omega}$, and line features detected by the laser scanner, ${}^S\boldsymbol{\ell}$, in order to estimate the quaternion, ${}^S\bar{q}_G$, which expresses the orientation of the global frame, $\{G\}$, with respect to the sensing frame $\{S\}$. The orientation is passed through a low-pass filter to obtain, ψ, which is an estimate of the person's heading direction. The second layer consists of the 2D position EKF which utilizes the person's heading direction and velocity, v, along with corner features extracted from the laser scan data, ${}^S\mathbf{p}_L$, in order to estimate the person's position inside the building, ${}^G\mathbf{p}$.

when considering the variations in body types between users. Significant sensor calibration and harness adjustment may be required in order to use such a system. Additionally, a body-mounted sensor package will likely interfere with common tasks such as sitting in a chair, and may prevent certain articles of clothing from being comfortably worn (e.g., a jacket). In contrast to the body-mounted paradigm, we propose a portable sensing package mounted on a white cane, which is hand carried and does not interfere with the person's normal activities (see Fig. 1). In addition, incorporating the cane as part of the navigation aid is of critical importance since it allows the visually impaired user to physically touch the environment with a tool that they already know and trust (Giudice and Legge 2008).

Considering our proposed sensor placement and the requirement that the system must be portable, one can appreciate the stark difference between indoor human localization and mobile (wheeled) robot localization. When constructing an estimator for the pose of a mobile robot, measurements of its linear and rotational velocities are often available from its wheel encoders, its motion is typically assumed to be planar, and its extrinsic sensors (e.g., laser scanner, sonar, or camera) are rigidly connected to its chassis (Thrun et al. 2005). In our scenario the sensors measuring the person's motion move in 3D and are not rigidly connected to each other, which makes the task of combining their information significantly more challenging. Specifically, we utilize a 2D laser scanner and a 3-axis gyroscope mounted near the white-cane handle to provide attitude information about the cane, and a lightweight, foot-mounted pedometer to measure the person's walking speed (Hesch and Roumeliotis 2007). Ideally, the information from these sources should be fused in a single pose estimator. However, all three sensors move in 3D and the coordinate transformation from the pedometer to the laser/gyro is unknown and time varying. These factors render our problem significantly more challenging than planar robot localization, which necessitates the design of a new estimation framework.

In order to address the challenges of indoor human localization, we propose a novel two-layered 2.5D estimator (see Fig. 2). In the first layer, rotational velocity measurements from the 3-axis gyroscope are combined with relative orientation information inferred from the laser-scanner data to estimate the 3D attitude of the cane. The second layer incorporates corner features extracted from the laser-scan data, linear velocity measurements from the pedometer, and a filtered version of the cane's yaw to estimate the 2D position of the person with respect to a known building map. The map is comprised of the 2D coordinates of corners in the environment (i.e., wall intersections at hallway junctions), and can be constructed either from the building blue prints or by a mobile robot who has previously explored the area. Our estimator can accommodate both perfectly known maps and maps with uncertainty in the corner locations. We also assume that a portion of the building's structural planes (e.g., the floor, wall, and ceiling planes) lie perpendicular to each other. We leverage the perpendicular portions of the building as a 3D structural compass to refine the attitude estimate of the cane.

The remainder of the paper is organized as follows: Section II reviews the relevant literature on obstacle avoidance, navigation, and localization systems for the visually impaired. Our method for estimating the 3D attitude of the white cane is discussed in Section III-A. Section III-B details the low-pass filter which extracts the heading of the person from the attitude estimate of the cane. The 2D position filter for estimating the person's location is presented in Section III-C. A description of the hardware setup is given in Section IV-A and experimental results are provided in Section IV-B. Lastly, the conclusions and future work are discussed in Section V.

II. Related Work

Recent work has focused on developing hazard detection aids for the visually impaired (Whitney 1998). These employ sensors for obstacle avoidance such as laser pointers (Yuan and Manduchi 2005), and sonars on a wheelchair (Bell et al. 1994), on a robot connected at the tip of a white cane (Ulrich and Borenstein 2001; Shim and Yoon 2002), or as part of a travel

aid (Shoval et al. 1998; Ran et al. 2004). Cameras have also been suggested for object description in terms of size and color (Hub et al. 2004; Wong et al. 2003), for obstacle detection in conjunction with tactile feedback (Zelek et al. 2003), and for image-to-sound auditory representations (Meijer 1992). While these devices augment the perceptual abilities of a visually impaired person and reduce the probability of an accident due to collision with an undetected obstacle, they cannot be explicitly used as wayfinding aids without the development of appropriate algorithms for determining the person's pose with respect to the building frame of reference.

Mobile robot pose estimation has received significant attention starting with the introduction of the Simultaneous Localization and Mapping (SLAM) problem (Moutarlier and Chatila 1989; Smith et al. 1990; Leonard and Durrant-Whyte 1991; Thrun et al. 2005). Most SLAM algorithms are based on the Extended Kalman Filter (EKF) (Williams et al. 2002) or Monte Carlo Localization (MCL) (Dellaert et al. 1999; Montemerlo et al. 2002). These methods typically rely on key assumptions that: (i) the robot is equipped with wheel encoders to measure its linear and rotational velocities, (ii) the robot's trajectory is planar, and (iii) the robot's sensors are rigidly connected to its chassis. Since these assumptions are violated for human pose estimation, there are only few attempts to apply this knowledge to assist the visually impaired in their everyday navigation tasks. Instead most relevant efforts have focused on GPS-based outdoor navigation for humans, which cannot be used inside a building (Makino et al. 1996; Balachandran et al. 2003; Ceranka and Niedzwiecki 2003; Ran et al. 2004).

An approach to indoor wayfinding for the visually impaired is presented in (Kulyukin et al. 2004a,b). In this case, an autonomous robot is attached at the end of a leash as a substitute for a guide dog, and localizes using information from a network of Radio Frequency Identification (RFID) tags. One of the main limitations of this approach is that mobility is restricted to places that a mobile robot can reach. This rules out areas where stairs or steps are part of the spatial layout, and tight spaces such as inside an office or a crowded room. Additionally, the weight and volume of the robot, negatively affect its portability by a commuter. Furthermore, this method requires instrumentation of buildings with RFIDs which is costly and time consuming; this is also the case for similar ultrasound (Ran et al. 2004) and infrared (IR) (Ertan et al. 1998) based systems.

Another tag-based system is presented by Tjan et al. (2005), in which visual markers are used to determine the person's position. The core of the system is a handheld camera termed the "magic flashlight," which processes image data searching for digital signs in the environment. These signs encode a digital signature (a barcode deviant) on a retro-reflective surface, and must be placed in the environment during a setup procedure. After constructing a database of digital-sign locations and linking this database to a building-layout schematic, information can be sent to the user such as directions or a description of their surroundings.

Several dead-reckoning methods have also been presented, which seek to estimate the person's pose without external references. In (Yun et al. 2007), a human pose tracker is presented which uses accelerometers and a magnetometer. Double integration of accelerometer measurements to obtain position estimates results in errors which are quadratic in the distance traveled. After employing gait detection and drift error correction methods (Cavallo et al. 2005; Sagawa et al. 2000), the authors present results which show only linear accumulation of error. Unfortunately, using a magnetometer indoors to measure orientation is highly susceptible to interference from metal objects (e.g., cabinets, doors, and structural components) and electrical wiring. However, the main limitation is that dead-reckoning systems continuously accumulate error (even if the rate of increase is low), and this prevents them from being used over extended periods of time.

The main contribution of this paper is the design and analysis of a 2.5D pose estimator, for indoor human localization. The implemented device comprises a light-weight sensing package mounted on a white cane and a foot-mounted pedometer. What distinguishes our localization aid from the previous approaches is: (i) our system is light-weight, portable, and cane-based, allowing the person to maintain physical contact with the environment, (ii) our aid is unobtrusive and does not require the user to wear a vest or harness (iii) we do not require the sensors to be rigidly attached to each other or to strictly move in a single plane, and (iv) we do not rely on building instrumentation (e.g., RFID tags or digital signs) to localize.

III. Problem Formulation

The estimation algorithm described in this work consists of two layers. In the first layer, the 3D attitude of the white cane is estimated using rotational velocity measurements from a 3-axis gyroscope and orientation updates from a 2D laser scanner, which exploits the perpendicular building planes as a 3D structural compass. The heading direction of the person is computed by passing the cane's yaw estimate through a low-pass filter. In the second layer, the 2D position of the person is estimated using the person's heading direction from the low-pass filter and the person's walking speed which is measured by a foot-mounted pedometer. The person's position is updated using laser-scanner observations of corner features in the environment (i.e., wall intersections at hallway junctions), which are matched to known corners in a building map (see Fig. 2). It is important to note, that while data from the laser scanner is utilized in both layers of the filter, we do not reuse information. Instead, we avoid statistical correlation by processing the even-indexed points in the first layer and the odd-indexed points in the second layer. This issue is discussed in more detail in Appendix A. We now turn our attention to the 3D attitude filter, and analytically study the observability properties of the linearized dynamical system. Subsequently, we introduce the low-pass filter, and the 2D position estimator for the person.

A. Layer 1: White-Cane Attitude Estimation

Our goal is to estimate the 3D attitude of sensing frame of reference, $\{S\}$ attached to the 3-axis gyroscope, which is

rotating with respect to the global frame of reference, $\{G\}$ fixed at the origin of the building map. This task requires us to first select an appropriate parameterization of the 3D orientation. Arguably the most popular method for specifying an orientation in 3D is by means of roll, pitch, and yaw (rpy) angles. While the rpy representation is elegant for its intuitive interpretation and ease of visualization, it suffers from singularities which, if not dealt with properly, might cause the estimate to become unstable. A more robust attitude representation is the quaternion of rotation

$$\bar{q} = {}^{S}\bar{q}_{G} = \begin{bmatrix}\mathbf{v}\,\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right)\end{bmatrix}^{T} \tag{1}$$

which expresses the orientation of $\{G\}$ with respect to $\{S\}$. The vector $\mathbf{v}$ is the unit vector along the axis of rotation, and θ is the rotation angle[1]. This representation of attitude is ideal because it is compact and singularity-free. For clarity, the notation employed in this paper results in "natural order" quaternion multiplication. As such, the symbol $\otimes$ denotes multiplication fulfilling ${}^{L_1}\bar{q}_{L_3} = {}^{L_1}\bar{q}_{L_2} \otimes {}^{L_2}\bar{q}_{L_3}$, which is the attitude rotation between successive frames (Breckenridge 1999).

Attitude estimation is accomplished through the use of an Extended Kalman Filter (EKF) which fuses measurements from proprioceptive and exteroceptive sensing devices. Specifically, rotational velocity measurements from a 3-axis gyroscope are integrated to track the attitude, and straight-line segments, corresponding to the intersection of the laser scan plane and perpendicular planes in the building, are utilized to update the computed estimate.

1) Attitude Propagation: The state vector $\mathbf{x}$ consists of the quaternion $\bar{q}$ and the gyroscope bias $\mathbf{b}$. The error state $\tilde{\mathbf{x}}$ comprises the attitude angle-error vector $\delta\boldsymbol{\theta}$ and the gyroscope bias error $\tilde{\mathbf{b}} = \mathbf{b} - \hat{\mathbf{b}}$, i.e.,

$$\mathbf{x} = \begin{bmatrix}\bar{q}\\ \mathbf{b}\end{bmatrix}, \quad \tilde{\mathbf{x}} = \begin{bmatrix}\delta\boldsymbol{\theta}\\ \tilde{\mathbf{b}}\end{bmatrix}. \tag{2}$$

Note that while the state vector $\mathbf{x}$ is 7×1, the error state $\tilde{\mathbf{x}}$ is 6×1. Many EKF formulations maintain equal sized state and error-state vectors, however, the quaternion of rotation is defined to have unit length; which would cause the corresponding covariance matrix to lose rank. To account for this, we utilize the attitude angle-error vector $\delta\boldsymbol{\theta}$ in the error state defined from the following relation:

$$\delta\bar{q} = \bar{q} \otimes \hat{\bar{q}}^{-1} \simeq \begin{bmatrix}\frac{1}{2}\delta\boldsymbol{\theta}^{T} & 1\end{bmatrix}^{T}. \tag{3}$$

The error quaternion $\delta\bar{q}$ denotes a small rotational error between the true, $\bar{q}$, and the estimated, $\hat{\bar{q}}$, attitude of the cane (Trawny and Roumeliotis 2005).

[1]Throughout this paper, ${}^{X}\mathbf{y}$ denotes the expression of a vector $\mathbf{y}$ with respect to frame $\{X\}$, ${}^{X}\mathbf{C}_{W}$ is the rotation matrix rotating vectors from frame $\{W\}$ to frame $\{X\}$, while ${}^{X}\bar{q}_{W}$ is the corresponding quaternion of rotation, and ${}^{X}\mathbf{p}_{W}$ is the position of the origin of frame $\{W\}$, expressed with respect to frame $\{X\}$. Additionally, $\mathbf{I}_m$ and $\mathbf{0}_m$ denote the $m\times m$ identity, and zero matrices, respectively. Finally, the symbol "$\hat{}$" denotes estimates while "$\tilde{}$" refers to error quantities, and we indicate measured quantities with the subscript "m".

a) Continuous-time system model: The time evolution of the attitude quaternion is described by the following first-order differential equation:

$$\dot{\bar{q}}(t) = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega})\,\bar{q}(t) \tag{4}$$

where $\boldsymbol{\omega}$ is the instantaneous rotational velocity and

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

The gyroscopes measure the rotational velocity of the cane, $\boldsymbol{\omega}_m$, expressed with respect to $\{S\}$, corrupted by sensor bias $\mathbf{b}$, as well as measurement noise $\mathbf{n}_r$, i.e.,

$$\boldsymbol{\omega}_m = \boldsymbol{\omega} + \mathbf{b} + \mathbf{n}_r \tag{5}$$

where the turn-rate noise $\mathbf{n}_r$ is distributed as a zero-mean white Gaussian process of strength $\sigma_r^2\mathbf{I}_3$. The sensor bias $\mathbf{b}$ is modeled as a random walk with

$$\dot{\mathbf{b}} = \mathbf{n}_w \tag{6}$$

where $\mathbf{n}_w$ is distributed as a zero-mean white Gaussian noise process of strength $\sigma_w^2\mathbf{I}_3$. Note that σ_r and σ_w are computed off-line using a standard gyroscope calibration procedure (IEEE Std. 647-2006 2006). Linearizing (4) and using (3), (5), and (6) yields the following error-state propagation equation

$$\begin{bmatrix}\dot{\delta\boldsymbol{\theta}}\\ \dot{\tilde{\mathbf{b}}}\end{bmatrix} = \begin{bmatrix}-\lfloor\hat{\boldsymbol{\omega}}\times\rfloor & -\mathbf{I}_3\\ \mathbf{0}_3 & \mathbf{0}_3\end{bmatrix}\begin{bmatrix}\delta\boldsymbol{\theta}\\ \tilde{\mathbf{b}}\end{bmatrix} + \begin{bmatrix}-\mathbf{I}_3 & \mathbf{0}_3\\ \mathbf{0}_3 & \mathbf{I}_3\end{bmatrix}\begin{bmatrix}\mathbf{n}_r\\ \mathbf{n}_w\end{bmatrix}$$
$$\dot{\tilde{\mathbf{x}}} = \mathbf{F}_c\,\tilde{\mathbf{x}} + \mathbf{G}_c\,\mathbf{n} \tag{7}$$

where $\hat{\boldsymbol{\omega}} = \boldsymbol{\omega}_m - \hat{\mathbf{b}}$.

b) Discrete-time implementation: During each propagation step of duration δt, the bias estimate is constant [see (6)]

$$\hat{\mathbf{b}}_{k+1|k} = \hat{\mathbf{b}}_{k|k} \tag{8}$$

and the quaternion estimate is propagated by integrating (4)

$$\hat{\bar{q}}_{k+1|k} = \begin{bmatrix}\frac{\hat{\boldsymbol{\omega}}_{k|k}}{||\hat{\boldsymbol{\omega}}_{k|k}||}\cdot\sin\left(\frac{||\hat{\boldsymbol{\omega}}_{k|k}||}{2}\delta t\right)\\ \cos\left(\frac{||\hat{\boldsymbol{\omega}}_{k|k}||}{2}\delta t\right)\end{bmatrix} \otimes \hat{\bar{q}}_{k|k} \tag{9}$$

where $||\hat{\boldsymbol{\omega}}_{k|k}|| = \sqrt{\hat{\boldsymbol{\omega}}_{k|k}^{T}\hat{\boldsymbol{\omega}}_{k|k}}$ and $\hat{\boldsymbol{\omega}}_{k|k} = \boldsymbol{\omega}_{m_k} - \hat{\mathbf{b}}_{k|k}$. The error-state covariance matrix is propagated as

$$\mathbf{P}_{k+1|k} = \boldsymbol{\Phi}_k\mathbf{P}_{k|k}\boldsymbol{\Phi}_k^{T} + \mathbf{Q}_{d_k} \tag{10}$$

where the discrete-time state transition matrix is

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

and the discrete-time noise covariance matrix

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

where $\mathbf{Q}_c = \begin{bmatrix}\sigma_r^2\mathbf{I}_3 & \mathbf{0}_3\\ \mathbf{0}_3 & \sigma_w^2\mathbf{I}_3\end{bmatrix}$. Both $\mathbf{Q}_{d_k}$ and $\boldsymbol{\Phi}_k$ are integrated in closed form and their detailed expressions are available in (Trawny and Roumeliotis 2005).

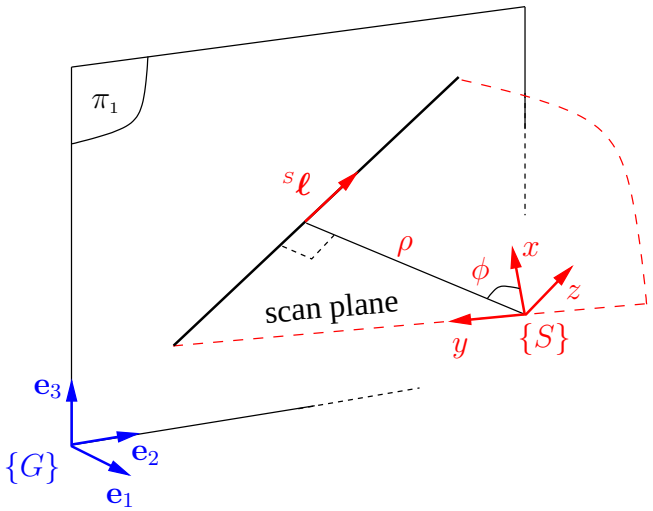

Fig. 3. ${}^S\boldsymbol{\ell}$ is the unit-vector direction of the line of intersection between the laser-scan plane and π_1, a structural plane of the building. Plane π_1 is of type Π_1 and is perpendicular to the unit-vector $\mathbf{e}_1$ which is one of the three which span the 3D space of the global frame $\{G\}$.

2) Attitude Update: The attitude propagation step allows us to track the time evolution of the orientation, however, due to the integration of gyroscope noise and bias, the state estimate continuously accumulates error. For this reason, it is necessary to update the attitude estimate with external information, which comes from the 2D laser scanner. During typical white cane usage, a walking person sweeps the tip of the cane in an arc on the floor in front of them in order to detect obstacles. This swinging motion causes the 2D laser-scan plane to intersect with different structural elements in the building, such as the floor, walls and ceiling (see Fig. 3). In almost all buildings, at least some portion of the structural elements are planar and lie perpendicular to each other. We exploit this perpendicularity as a 3D structural compass to update the attitude estimate, however, we do not require all walls to be planar, or even all planes to be perpendicular. As we will show, each laser-scan measurement must first be matched to one of the principal map planes before it is utilized to update the orientation, hence, measurements corresponding to non-perpendicular planes, or to non-planar objects in the environment are simply discarded.

The intersection between the laser-scan plane, and structural planes in the building appear as straight-line segments in the laser-scan data. We employ weighted line fitting (Pfister et al. 2003) on the even-indexed data points to estimate (ρ, ϕ) which are the polar parameters of the line. This line-fitting algorithm robustly detects lines in the raw laser-scan data, while rejecting points which correspond to intersections with non-planar obstacles in the environment (refer to Appendix A for a complete discussion). The direction of each measured line provides information about the relative orientation between frames $\{S\}$ and $\{G\}$, and we exploit this information to update the attitude estimate.

Inside the building, $\{G\}$ is fixed at the origin of the building map and its principal axes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ are orthogonal to the building's prominent structural planes (i.e., the walls, ceiling, and floor that lie perpendicular to each other). We categorize each structural plane according to which principal axis it is orthogonal, such that the three plane types are Π_i, $i = 1, \ldots, 3$:

$$\mathbf{e}_1 = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^T \perp \Pi_1, \quad \mathbf{e}_2 = \begin{bmatrix}0 & 1 & 0\end{bmatrix}^T \perp \Pi_2$$
$$\mathbf{e}_3 = \begin{bmatrix}0 & 0 & 1\end{bmatrix}^T \perp \Pi_3.$$

To simplify the discussion, we focus on a single measurement of a plane, π_i of type Π_i, whose normal vector is $\mathbf{e}_i$. The unit-vector ${}^G\boldsymbol{\ell}$ denotes the direction of the line of intersection between the laser-scan plane and π_i, expressed with respect to $\{G\}$ (see Fig. 3). Since, ${}^G\boldsymbol{\ell}$ lies in π_i and $\mathbf{e}_i$ is orthogonal to it, their inner product should be zero, i.e.,

$$\mathbf{e}_i^T {}^G\boldsymbol{\ell} = 0, \quad i \in \{1, 2, 3\}. \tag{11}$$

The measurement constraint is derived by rewriting (11) using the transformation, ${}^G\boldsymbol{\ell} = C^T(\bar{q})\,{}^S\boldsymbol{\ell}$, i.e.,

$$z = \mathbf{e}_i^T C^T(\bar{q})\,{}^S\boldsymbol{\ell} = 0 \tag{12}$$

where the rotational matrix $C^T(\bar{q})$ rotates vectors expressed with respect to $\{S\}$ into $\{G\}$. Since ${}^S\boldsymbol{\ell}$ is the unit-vector direction of a line in the x-y plane of the sensor frame, it can be expressed as ${}^S\boldsymbol{\ell} = \begin{bmatrix}\sin(\phi) & -\cos(\phi) & 0\end{bmatrix}^T$ where ϕ is the angle from the sensor's x-axis to the perpendicular to the line direction.

Applying the expectation operator on the first-order Taylor series approximation of (12), we obtain the expected measurement equation

$$\hat{z} = \mathbf{e}_i^T C^T\left(\hat{\bar{q}}\right) {}^S\boldsymbol{\ell}_m \tag{13}$$

where ${}^S\boldsymbol{\ell}_m = \begin{bmatrix}\sin(\phi_m) & -\cos(\phi_m) & 0\end{bmatrix}$ denotes the measured line direction and $\phi_m = \phi - \tilde{\phi}$ is the measured angle to the perpendicular to the line direction. The angle error, $\tilde{\phi}$, accounts for measurement noise as well as line-fitting inaccuracies, and is modeled as zero-mean white Gaussian noise with variance σ_ϕ^2. Employing (12) and (13), the measurement error is:

$$\begin{aligned}\tilde{z} \simeq\ & \begin{bmatrix}-\mathbf{e}_i^T C^T\left(\hat{\bar{q}}\right) \lfloor {}^S\boldsymbol{\ell}_m \times \rfloor & \mathbf{0}_{1\times 3}\end{bmatrix} \begin{bmatrix}\delta\boldsymbol{\theta} \\ \tilde{\mathbf{b}}\end{bmatrix} \\ & + \mathbf{e}_i^T C^T\left(\hat{\bar{q}}\right) \lfloor \mathbf{e_3} \times \rfloor {}^S\boldsymbol{\ell}_m \tilde{\phi} \\ =\ & \mathbf{h}^T \tilde{\mathbf{x}} + \gamma \tilde{\phi}\end{aligned} \tag{14}$$

where $\mathbf{h}^T$ and γ are the measurement Jacobians with respect to $\mathbf{x}$ and ϕ, respectively. We also compute the measurement residual as $r = z - \hat{z} = -\mathbf{e}_i^T C^T\left(\hat{\bar{q}}\right) {}^S\boldsymbol{\ell}_m$.

Prior to applying an attitude update, we first match the measured line direction to one of the principal plane directions in the building. To do so, we compute the Mahalanobis distance for the measurement matching each building direction, i.e.,

$$d_i = \frac{r_i^2}{\mathbf{h}^T \mathbf{P}_{k+1|k} \mathbf{h} + \sigma_\phi^2 \gamma^2}, \quad i = 1, \ldots, 3. \tag{15}$$

If the minimum distance is less than a probabilistic threshold, then the measurement is matched to the corresponding plane and we proceed with the attitude update; otherwise we discard

the line measurement. Hence, the filter is robust to erroneously detected lines in the laser data.

In order to update the attitude estimate, we first compute the Kalman gain:

$$\mathbf{k} = \frac{1}{\mathbf{h}^T\mathbf{P}_{k+1|k}\mathbf{h} + \sigma_\phi^2\gamma^2}\mathbf{P}_{k+1|k}\mathbf{h}. \quad (16)$$

The computed state correction is

$$\begin{bmatrix}\delta\boldsymbol{\theta}_{corr}\\ \mathbf{b}_{corr}\end{bmatrix} = \mathbf{k}r, \quad (17)$$

where $\delta\boldsymbol{\theta}_{corr}$ is the correction to be applied to the orientation estimate and $\mathbf{b}_{corr}$ is the correction to be applied to the bias estimate. From $\delta\boldsymbol{\theta}_{corr}$ the error quaternion estimate is computed as

$$\delta\hat{\bar{q}} = \frac{1}{\sqrt{1+\frac{1}{4}\delta\boldsymbol{\theta}_{corr}^T\delta\boldsymbol{\theta}_{corr}}}\begin{bmatrix}\frac{1}{2}\delta\boldsymbol{\theta}_{corr}\\ 1\end{bmatrix}. \quad (18)$$

The quaternion and the bias estimates are updated as

$$\hat{\bar{q}}_{k+1|k+1} = \delta\hat{\bar{q}} \otimes \hat{\bar{q}}_{k+1|k}, \quad (19)$$

$$\hat{\mathbf{b}}_{k+1|k+1} = \hat{\mathbf{b}}_{k+1|k} + \mathbf{b}_{corr}. \quad (20)$$

The last step in this process is the covariance update

$$\mathbf{P}_{k+1|k+1} = (\mathbf{I}_6 - \mathbf{k}\mathbf{h}^T)\mathbf{P}_{k+1|k}(\mathbf{I}_6 - \mathbf{k}\mathbf{h}^T)^T + \sigma_\phi^2\gamma^2\mathbf{k}\mathbf{k}^T. \quad (21)$$

3) Attitude Observability: Due to physical limitations of planar laser scanning, we can only resolve all three degrees of rotational freedom from a stationary vantage point in a few scenarios (Chen 1991). The attitude of the cane, however, is observable when the cane is in motion. In what follows (see Lemma 1), we prove that the observability requirements are satisfied while the white cane is being used.

Lemma 1: All three degrees of rotational freedom are observable when the laser scanner measures at least two main directions in space, at different time steps, over a given interval.

Proof: Since observability of the attitude entails observability of the bias, we focus on a simplified measurement model with only the quaternion as state variable, and measurement matrix [see (14)]

$$\mathbf{h}_i^T(t_k) = -\mathbf{e}_i^T C^T(\bar{q}(t_k))\,\lfloor {}^S\boldsymbol{\ell}_i(t_k)\times\rfloor. \quad (22)$$

In order to establish observability of the attitude, it suffices to show that the observability Gramian

$$\mathbf{M} = \sum_{k=1}^{N}\sum_i \boldsymbol{\Phi}^T(t_k,0)\mathbf{h}_i(t_k)\mathbf{h}_i^T(t_k)\boldsymbol{\Phi}(t_k,0) \quad (23)$$

is of full rank for some finite N (Maybeck 1979). Noting that in this case the state transition matrix $\boldsymbol{\Phi}(t_k,0) = -\int_0^{t_k}\lfloor\boldsymbol{\omega}\times\rfloor d\tau = C(\bar{q}(t_k))$ (Trawny and Roumeliotis 2005), the observability Gramian can be written as the sum of the following vector outer products

$$\mathbf{M} = \sum_{k=1}^{N}\sum_i {}^G\mathbf{h}_i(t_k){}^G\mathbf{h}_i^T(t_k)$$

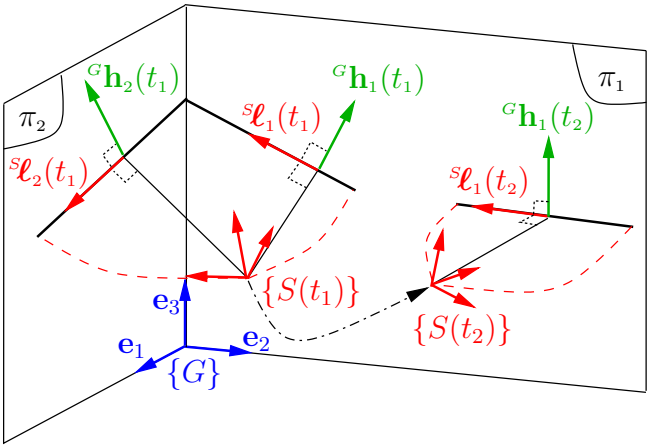

Fig. 4. An illustration of the observability property being fulfilled. In this instance, planes π_1 and π_2 are observed in the sensor frame $\{S\}$ at time t_1 denoted $\{S(t_1)\}$, subsequently, π_1 is measured at time t_2. The corresponding vectors ${}^G\mathbf{h}_1(t_1)$, ${}^G\mathbf{h}_2(t_1)$, and ${}^G\mathbf{h}_1(t_2)$ span the 3D space, hence the observability Gramian [see (23)] is full rank.

where ${}^G\mathbf{h}_i(t_k) = {}^G\boldsymbol{\ell}_i(t_k)\times\mathbf{e}_i$. If over a period of time the sensor observes two or more surfaces with normals $\mathbf{e}_i$ that span $\mathbb{R}^2$, and recalling that ${}^G\boldsymbol{\ell}_i^T\mathbf{e}_i = 0$ [see (11)], then the vectors ${}^G\mathbf{h}_i$ span the 3D space, ensuring $\mathbf{M}$ to be of full rank. Fig. 4 depicts a scenario in which the observability property is fulfilled. ■

B. Person Heading Estimation

In the previous section, we introduced a method to estimate the 3D attitude of the cane with a 3-axis gyroscope, while utilizing the laser-scanner observations of the perpendicular building planes as a 3D structural compass to update the attitude. In what follows, we describe how to extract the person's heading direction from the 3D attitude estimate of the cane.

During typical usage, a person walking with a white cane will hold the cane out in front of them with its tip on the ground, swinging it to the right and to the left in order to detect potential obstacles and hazards in their walking path. The yaw angle of the cane, which we obtain by converting the quaternion estimate to rpy, will be a cyclostationary random process whose amplitude, phase, and frequency may all change (see Fig. 5). Although this random process exhibits variability depending on the person and trajectory, we have found that the heading of the person can be well approximated as the mean value of the cane's yaw over a period of swinging. Thus, we can apply signal-processing techniques to obtain the person's heading direction as the dc component of the cane's yaw estimate. This motivates the use of a Finite Impulse Response (FIR) filter, to remove the high-frequency component of the yaw. The attitude is propagated at 100 Hz, and we have observed over several trials with different users that the high frequency component of the cane's yaw signal does not fall below 0.5 Hz. In order to reduce the number of taps needed by the FIR filter, the yaw signal is down-sampled by a factor of 40. The signal is then filtered using a 7^{th} order FIR filter with Kaiser window $\beta = 0.5$ and normalized cut-off frequency

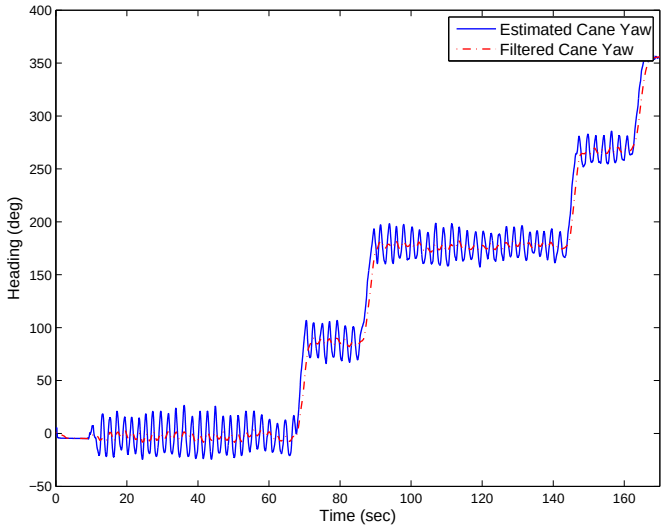

Fig. 5. The yaw component of the cane's attitude estimate plotted along with the low-pass filtered yaw. In the results presented here, the frequency of the cane's swinging motion is between 2 and 2.5 Hz, and the amplitude is between 20 and 30 deg.

0.02. Fig. 5 depicts the yaw component of the cane's attitude estimate along with the FIR filter's output.

There are several important considerations which need to be addressed when employing this procedure. The first is how close the mean of the yaw signal needs to match the person's heading direction. We have seen over several trials with different people that if the dc component is within two or three degrees of the true heading direction, then the filter performs well. In order to accommodate users who exhibit significant bias (e.g., ten degrees or more), we are investigating methods to detect and estimate this bias on-line. The second important consideration is the time-delay introduced by the FIR filter. Although the filter has only seven coefficient-delay pairs, there is a delay of 1.2 sec because the down-sampled yaw signal has a sampling frequency of 2.5 Hz. As a result, an equivalent time-lag exists in the position estimate. However, since the average walking speed of a person is between 1.25 and 1.51 m/sec (Knoblauch et al. 1996), this time lag does not prohibit real-time operation. A third consideration is how the heading estimate is affected when the cane collides with obstacles in the environment. Fortunately, the gyroscope sampling frequency is high (100 Hz), and it accurately captures quick changes in the cane's rotational velocity. Thus, even during collisions the cane's 3D attitude is estimated correctly, and the person's heading direction is reliably computed.

C. Layer 2: Person Position Estimation

Estimating the position of a person within a building is treated as a 2D estimation problem in which each level of the building is a separate planar environment. In this scenario, we utilize a building map comprised of corner features that correspond to the wall intersections at hallway junctions. We assume the map is available *a priori*, and is computed either from the building blueprints or by a robot which has previously explored the environment. The map may be perfectly known or may have some uncertainty, and we accommodate both cases in our estimator. While a person is traversing a single floor, their motion will be constrained on the plane of that floor. This allows for the use of a 2D odometry propagation model.

1) Position Propagation: The non-holonomic formulation of the odometry state equations requires that linear and rotational velocity measurements be available during the propagation stage. These constraints are relaxed for the case of a person. Linear velocity measurements are provided by a foot-mounted wireless pedometer. Note that rotational velocity measurements need not be accessible as the person's heading direction is available from the FIR filter. This odometric model is also known as direct heading odometry (Kelly 2004). The state and state-estimate propagation equations are:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} + v\delta t \begin{bmatrix} \cos(\psi) \\ \sin(\psi) \end{bmatrix} \tag{24}$$

$$\begin{bmatrix} \hat{x}_{k+1} \\ \hat{y}_{k+1} \end{bmatrix} = \begin{bmatrix} \hat{x}_k \\ \hat{y}_k \end{bmatrix} + v_m\delta t \begin{bmatrix} \cos(\psi_m) \\ \sin(\psi_m) \end{bmatrix} \tag{25}$$

where $\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix}^T$ and ψ are the position and heading of the person, and v is the average velocity during the time interval δt. The velocity measured by the pedometer and the measured heading direction are

$$v_m = v + n_v \ , \quad \psi_m = \psi + n_\psi$$

where the errors in velocity, n_v, and heading, n_ψ, are zero-mean white Gaussian processes with variances σ_v^2 and σ_ψ^2, respectively. The linearized error-state model is

$$\begin{bmatrix} \tilde{x}_{k+1} \\ \tilde{y}_{k+1} \end{bmatrix} = \begin{bmatrix} \tilde{x}_k \\ \tilde{y}_k \end{bmatrix} + \begin{bmatrix} -\delta t\cos(\psi_m) & v_m\delta t\sin(\psi_m) \\ -\delta t\sin(\psi_m) & -v_m\delta t\cos(\psi_m) \end{bmatrix} \begin{bmatrix} n_v \\ n_\psi \end{bmatrix}$$
$$\tilde{\mathbf{x}}_{k+1} = \tilde{\mathbf{x}}_k + \mathbf{\Gamma n} \tag{26}$$

and thus the covariance propagation is computed as

$$\mathbf{P}_{k+1|k} = \mathbf{P}_{k|k} + \mathbf{\Gamma Q \Gamma}^T, \ \text{where } \mathbf{Q} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_\psi^2 \end{bmatrix}. \tag{27}$$

2) Position Update: The person's position estimate is updated by incorporating relative position measurements to known landmarks in the environment. Although the selection of features is flexible, using corners at hallway intersections is a good choice for an indoor environment because they are prevalent and can be extracted reliably from the laser scanner data (see Appendix A). A corner is identified as the intersection between two lines extracted from the odd-indexed laser data. Since we are only considering corners at the intersection of structural planes (i.e., wall intersections at hallway junctions) each line must first be matched to one of the building planes [see (15)]. The end points of the two lines must be closer than a distance threshold (e.g., 5 cm), which ensures that the two lines truly intersect at the corner. In this way, our corner extraction method is robust to clutter in the environment, and we reliably match measured corners to corners in the map using a probabilistic data association step. The robustness and reliability of the line fitting and corner extraction procedures are discussed further in Appendix A, and for now we turn our attention to the mathematical model used for the position update procedure.

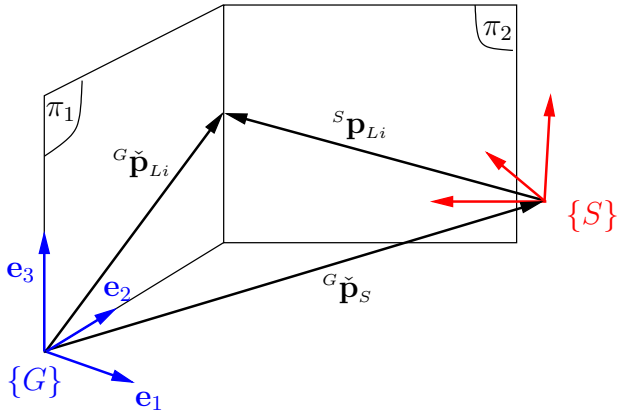

Fig. 6. The corner measurement: ${}^S\mathbf{p}_{Li}$ is the point of intersection between the corner and the laser scan plane, expressed in $\{S\}$. The vector ${}^G\check{\mathbf{p}}_S$ is the origin of $\{S\}$ with respect to $\{G\}$, and ${}^G\check{\mathbf{p}}_{Li}$ is the corner with respect to $\{G\}$.

The relative position measurement is derived using the 3D geometric relationship depicted in Fig. 6, i.e.,

$$ {}^G\check{\mathbf{p}}_{Li} = {}^G\check{\mathbf{p}}_S + C^T(\bar{q})\,{}^S\mathbf{p}_{Li} \tag{28} $$

where ${}^S\mathbf{p}_{Li}$ is the 3D vector describing the intersection of the corner L_i with the laser scan plane, ${}^G\check{\mathbf{p}}_{Li}$ is the 3D vector from the origin of $\{G\}$ to the intersection point, and ${}^G\check{\mathbf{p}}_S$ is the 3D position of the sensor with respect to $\{G\}$. The 2D position of the person $\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix}^T$, is approximated by the projection of the sensor's position in the x-y plane of $\{G\}$. Hence, the first two components of ${}^G\check{\mathbf{p}}_S = \begin{bmatrix} x & y & z \end{bmatrix}^T$ comprise the state vector of the position filter [see (24)].

The measurement model used for updating the person's position is obtained by bringing all quantities in (28) to the left-hand side and applying the 2×3 projection matrix $\mathbf{T} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix}^T$ on both sides[2], i.e.,

$$ \begin{aligned} \mathbf{z} &= \mathbf{T}\left({}^G\check{\mathbf{p}}_{Li} - {}^G\check{\mathbf{p}}_S - C^T(\bar{q})\,{}^S\mathbf{p}_{Li}\right) \\ &= {}^G\mathbf{p}_{Li} - {}^G\mathbf{p}_S - \mathbf{T}\,C^T(\bar{q})\,{}^S\mathbf{p}_{Li} = \mathbf{0}_{2\times 1} \end{aligned} \tag{29} $$

where ${}^G\mathbf{p}_{Li}$ is the 2D position of the landmark L_i which is known from the map, and ${}^G\mathbf{p}_S = \begin{bmatrix} x & y \end{bmatrix}^T = \mathbf{x}$ is the 2D position of the sensor in $\{G\}$ (or equivalently, the 2D position of the person). Applying the expectation operator on the first-order Taylor series approximation of (29), we compute the expected measurement as

$$ \hat{\mathbf{z}} = {}^G\mathbf{p}_{Li} - {}^G\hat{\mathbf{p}}_S - \mathbf{T}\,C^T\left(\hat{\bar{q}}\right){}^S\mathbf{p}_{Lmi} \tag{30} $$

where ${}^G\hat{\mathbf{p}}_S$ is the estimated 2D position of the origin of $\{S\}$ with respect to $\{G\}$, while ${}^S\mathbf{p}_{Lmi} = {}^S\mathbf{p}_{Li} - {}^S\tilde{\mathbf{p}}_{Li}$ is the measured 3D position of the intersection of the laser scan plane with corner L_i expressed in $\{S\}$. Finally, differentiation of (29), provides the linearized error model

$$ \begin{aligned} \tilde{\mathbf{z}} &\simeq -{}^G\tilde{\mathbf{p}}_S + \mathbf{T}\,C^T\left(\hat{\bar{q}}\right)\lfloor {}^S\mathbf{p}_{Lmi}\times\rfloor\delta\boldsymbol{\theta} - \mathbf{T}\,C^T\left(\hat{\bar{q}}\right){}^S\tilde{\mathbf{p}}_{Li} \\ &= -\mathbf{I}_2\,\tilde{\mathbf{x}} + \begin{bmatrix} \mathbf{T}\,C^T\left(\hat{\bar{q}}\right)\lfloor {}^S\mathbf{p}_{Lmi}\times\rfloor & -\mathbf{T}\,C^T\left(\hat{\bar{q}}\right) \end{bmatrix}\begin{bmatrix} \delta\boldsymbol{\theta} \\ {}^S\tilde{\mathbf{p}}_{Li} \end{bmatrix} \\ &= \mathbf{H}\,\tilde{\mathbf{x}} + \mathbf{G}\,\mathbf{n}_p \end{aligned} \tag{31} $$

where $\mathbf{H}$ and $\mathbf{G}$ are the Jacobians with respect to the state and measurement, respectively. The error vector $\mathbf{n}_p$ is comprised of the quaternion error vector, $\delta\boldsymbol{\theta}$, and the error in the corner measurement, ${}^S\tilde{\mathbf{p}}_{Li}$. Note that $\mathbf{n}_p$ is zero-mean, white Gaussian noise, with covariance $\mathbf{R}$.

In order to perform data association, we form the measurement residual, $\mathbf{r} = \mathbf{z} - \hat{\mathbf{z}} = -\hat{\mathbf{z}}$ [see (29)], and the covariance of the measurement residual $\mathbf{S} = \mathbf{H}\,\mathbf{P}_{k+1|k}\mathbf{H}^T + \mathbf{G}\,\mathbf{R}\,\mathbf{G}^T$. We compute the Mahalanobis distance for the measurement corresponding to each landmark, i.e., $d_i = \mathbf{r}^T\mathbf{S}^{-1}\mathbf{r}$, and we only declare a measurement-to-landmark match if the minimum distance is below a preselected probabilistic threshold. In this way, we ensure that the filter is robust to erroneously detected corners.

If the measurement is found to match a corner in the map, then we proceed with the update, and compute the Kalman gain

$$ \mathbf{K} = \mathbf{P}_{k+1|k}\mathbf{H}^T\left(\mathbf{H}\,\mathbf{P}_{k+1|k}\mathbf{H}^T + \mathbf{G}\,\mathbf{R}\,\mathbf{G}^T\right)^{-1}. \tag{32} $$

The state estimate is updated as

$$ \hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}\,\mathbf{r} \tag{33} $$

and lastly, the updated covariance is computed as

$$ \begin{aligned} \mathbf{P}_{k+1|k+1} = &\left(\mathbf{I}_2 - \mathbf{K}\,\mathbf{H}\right)\mathbf{P}_{k+1|k}\left(\mathbf{I}_2 - \mathbf{K}\,\mathbf{H}\right)^T \\ &+ \mathbf{K}\,\mathbf{G}\,\mathbf{R}\,\mathbf{G}^T\mathbf{K}^T. \end{aligned} \tag{34} $$

Our position update procedure readily extends to the case of an imperfect map. We define the landmark position error as ${}^G\tilde{\mathbf{p}}_{L_i} = {}^G\mathbf{p}_{L_i} - {}^G\hat{\mathbf{p}}_{L_i}$, where ${}^G\tilde{\mathbf{p}}_{L_i}$ is distributed as zero-mean white Gaussian noise with covariance $\mathbf{R}_{L_i}$. We modify the linearized-error model to account for the uncertainty in ${}^G\mathbf{p}_{L_i}$, i.e.,

$$ \begin{aligned} \tilde{\mathbf{z}} &\simeq \mathbf{H}\,\tilde{\mathbf{x}} + \begin{bmatrix} -\mathbf{I}_2 & \mathbf{G} \end{bmatrix}\begin{bmatrix} {}^G\tilde{\mathbf{p}}_{L_i} \\ \mathbf{n}_p \end{bmatrix} \\ &= \mathbf{H}\,\tilde{\mathbf{x}} + \mathbf{G}'\,\mathbf{n}'_p \end{aligned} \tag{35} $$

where $\mathbf{n}'_p$ and $\mathbf{G}'$ are the new measurement noise, and its Jacobian, respectively. We then compute (32), (33), and (34) using the new linearized error model, and perform the position update.

IV. Experimental Results

A. Hardware Description

When designing the sensor platform used in this work, the main criterion for the sensor selection and placement was that the electronics should be lightweight and unobtrusive to the user. For this reason two of the three primary sensors are affixed on the white cane (see Fig. 7), and the third is foot mounted. The sensor bay is attached near the handle of the cane which has length 1.27 m. The cane is a lightweight,

[2]Since (28) contains unknown quantities which are neither measured nor estimated (i.e., the z-coordinates of ${}^G\check{\mathbf{p}}_{Li}$ and ${}^G\check{\mathbf{p}}_S$), the 2D relative position constraint is obtained by projecting (28) onto the x-y plane of $\{G\}$.

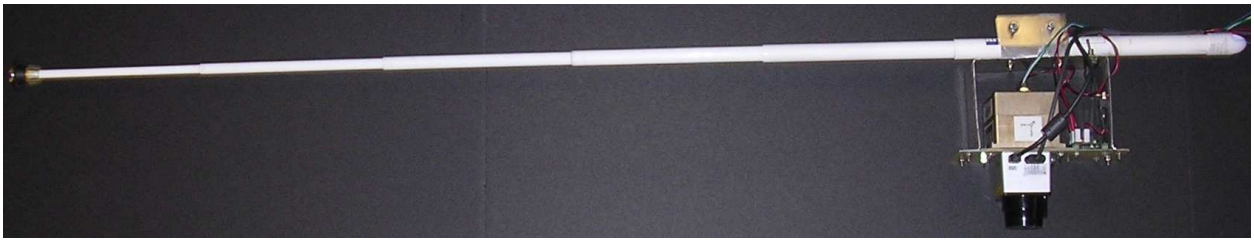

Fig. 7. A view of the white cane and sensing hardware. The sensor bay is mounted near the handle of the cane. Note that the tip of the cane is unobstructed, and the user maintains the ability to physically sense the world. The weight of the sensor package is approximately 550 g.

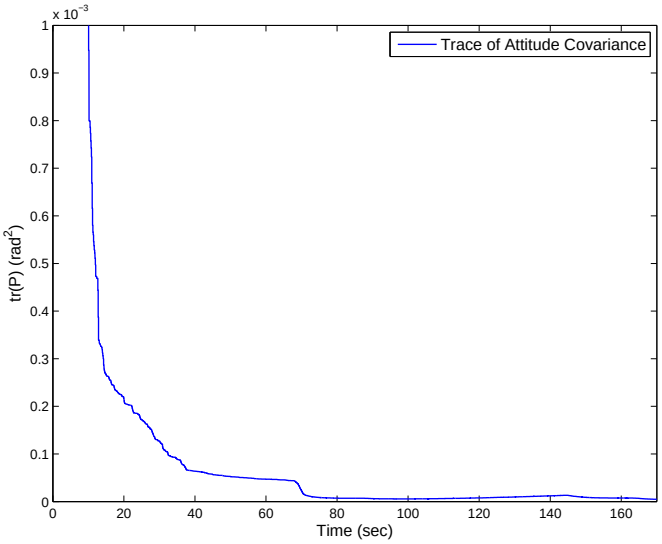

Fig. 8. The trace of the attitude covariance demonstrates that while the cane is stationary (first 10 sec during bias initialization) its attitude is initially unobservable, however, it becomes observable when the cane is in motion.

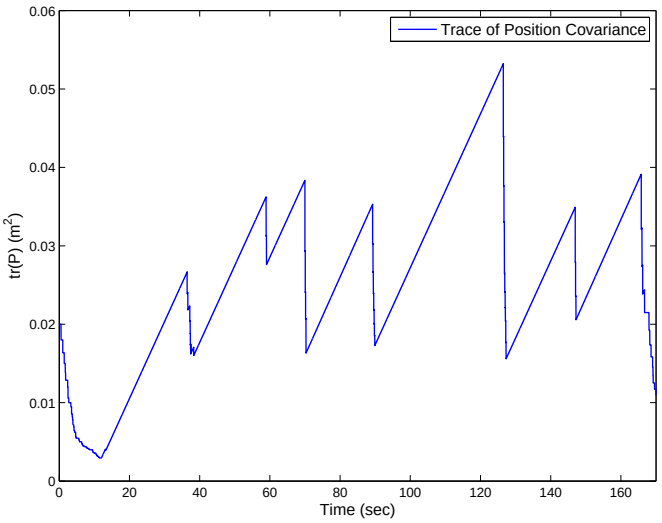

Fig. 9. The trace of the position covariance verifies that the positioning uncertainty remains bounded, but grows linearly with time between consecutive position updates.

carbon-fiber model, selected so that the total weight, including sensors, is approximately the same as a standard white cane.

The laser scanner is an URG-X002S which measures 5 cm by 5 cm by 7 cm. It has an angular scan range of 240 deg, with an accuracy of $\pm 1\%$ of the measurement for distances 1 m to 4 m. Closer than 1 m the measurement accuracy is ± 10 mm. The laser scanner can measure distances ranging from 0.02 m to 4 m. The scanner weighs 160 g and consumes 2.5 W at 5 V. The 3-axis gyroscope is an ISIS Inertial Measurement Unit (IMU), with an angular-rate range of ± 90 deg/sec. Over an RS-232 connection, the ISIS IMU provides measurements at 100 Hz. The weight of the sensor is 363 g, and the power consumption is 6.72 W at 12 V. The IMU measures 5.5 cm by 6.5 cm by 7 cm. The pedometer is 8.5 cm by 3.5 cm by 3.5 cm, and transmits communication packets via Bluetooth at a rate of 1 Hz. Power is supplied to the sensor bay by 8 RCR123A rechargeable lithium-ion batteries which are placed inside the cane handle and have a combined weight of 136 g. All the sensors were interfaced via USB, RS-232, and Bluetooth, respectively, to an ultra-portable Sony Vaio (VGN-UX series) which has dimensions 15 cm by 9.5 cm by 3.2 cm, and weighs 540 g. The computer can be carried in hand, or in a case connected to the user's belt. The real-time software components are written in C++, whereas the software for simulation and data plotting is written in Matlab.

B. Experimental Results

The estimation algorithm described in this paper was tested in an indoor environment on a closed loop of path length 130 m. Twenty-one corners along this loop were known *a priori* from the building blueprints, and were used as features for position updates as described in Section III-C2. While walking around, the user avoided potential hazards by swinging the cane to the right and to the left at a frequency of 2 to 2.5 Hz with amplitude of 20 to 30 deg. Fig. 10 shows the estimated trajectory super-imposed on the floor diagram. The striped regions in the figure depict obstacles such as couches and garbage cans, which are not detailed in the building blueprint. Additionally, some of the doors along the hallways were open, while others were closed. During testing there was a normal flow of pedestrian traffic through the hallways. All of the corners in the map are shown as boxes, and every measurement which was used to update the position estimate is marked with a line to the corresponding corner.

Results for the first layer (3D attitude EKF) are plotted in Fig. 8, which depicts the trace of the attitude covariance matrix with respect to time. During this experiment, the white cane was initially stationary for 10 sec for the purpose of gyroscope bias initialization. As evident, the trace of the attitude covariance becomes bounded once the cane is in motion. This supports the observability analysis of the system,

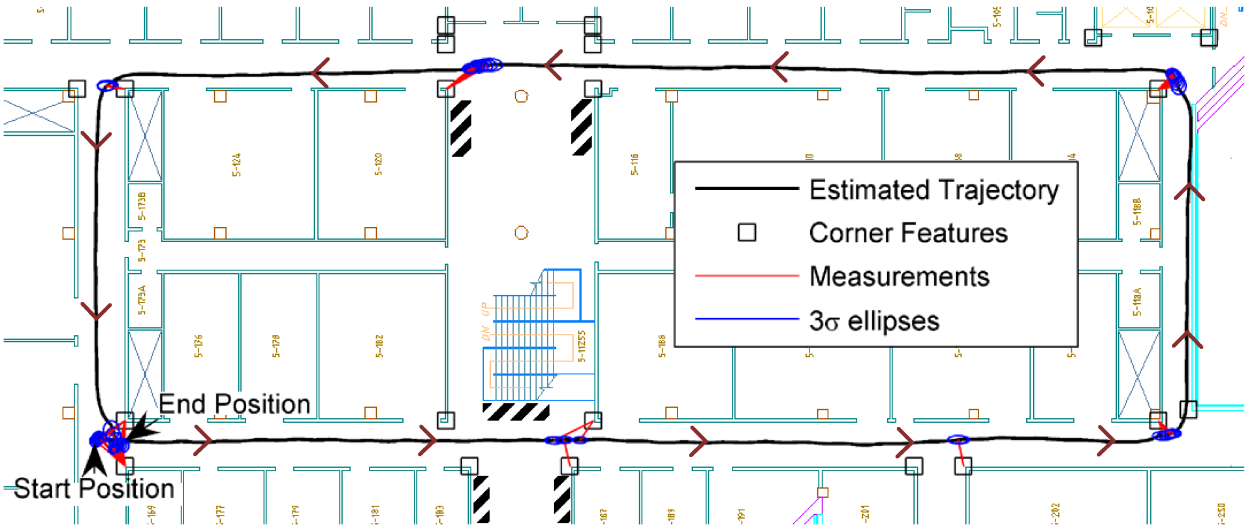

Fig. 10. The estimated trajectory is plotted (black), overlaid on the building blueprints, for an experimental trial on a 130 m loop. The start and end locations are in the lower left-hand corner of the map. The direction of travel was counter-clockwise around the loop, which is indicated by the arrows on the trajectory (brown). Corner measurements are depicted (red), along with the corresponding 3σ position uncertainty (blue) at the time when each measurement occurred. The squares (black) denote the 21 corners that were included in the *a priori* map. The striped regions on the figure depict locations of furniture and other objects not represented in the blueprint.

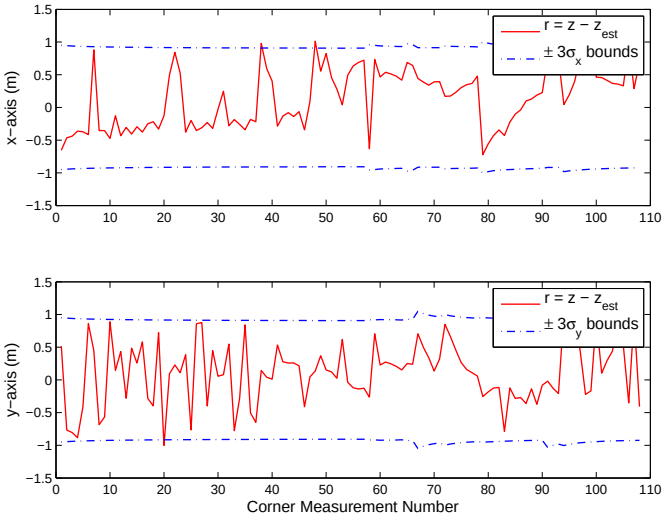

Fig. 11. The x- and y-axis position measurement residuals plotted with their corresponding 3σ bounds computed from the residual covariance. This is an indication of the consistency of the filter, and one of the key characteristics which we examine in real-world experiments when ground truth is not available.

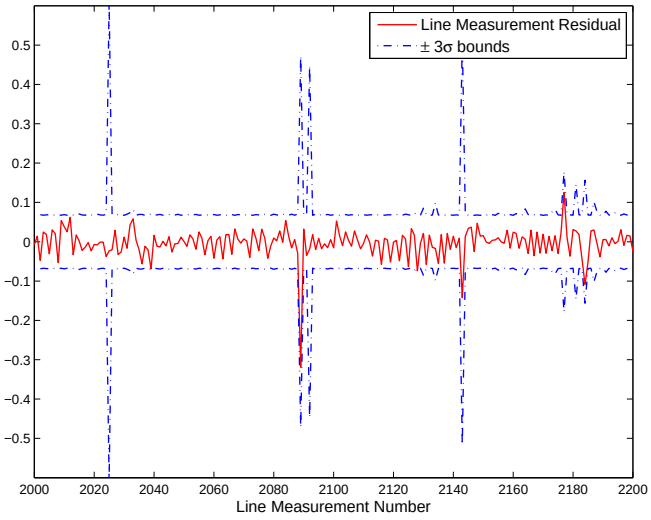

Fig. 12. The residuals of the attitude update measurements plotted with the 3σ bounds computed from the residual covariance. Note that the units of the vertical-axis are intentionally omitted considering that this measurement residual is by nature unit-less (difference of unit-vector dot-products). In order to clearly see the residuals and bounds we are showing a subset (approx. 200) of the $5,000$ attitude updates which were processed during this experiment.

since the laser scanner observes planes whose normal vectors span $\mathbb{R}^3$ while in motion.

The uncertainty in the position estimate is small (max. $\sigma = 0.16$ m), even though the number of position update measurements are few (only 9 of the corners were detected in approximately 110 laser scans). The reason the position filter is precise despite the relatively infrequent position update measurements is due to the high quality of the heading estimates provided by the attitude filter. The attitude estimates of the cane are accurate as a result of more than $5,000$ relative orientation measurements obtained during the experiment (see Fig. 8). Based on the analysis of (Mourikis and Roumeliotis 2005) we can infer that when the orientation error is bounded, i.e., $\sigma_\psi \leq \sigma_{\psi_0}$, the position covariance grows as

$$\mathbf{P}\,(t) \leq 0.5\left(\sigma_v^2 + \sigma_{\psi_0}^2 v^2 \delta t^2\right)\ t = \alpha\ t. \tag{36}$$

In our experiments, $\delta t = 0.1$ sec and $\alpha = 9.8204 \times 10^{-4}$. This means that for the case of direct heading odometry (Kelly 2004), the position uncertainty grows approximately linearly with time between consecutive position updates. Thus, even when detecting only a small number of corners, the position filter maintains a good estimate. This argument is corroborated by the time evolution of the trace of the position covariance. The value of the trace never exceeds 0.054 m^2 which corre-

sponds to approximately 0.16 m 1σ error in each direction (see Fig. 9). Furthermore, the consistency of both the attitude and position filters is verified. The x- and y-components of the position measurement residuals fall within the 3σ bounds of the residual covariance (see Fig. 11). Similarly, the orientation measurement residuals lie within the 3σ bounds of the residual covariance (see Fig. 12). The filter consistency is also verified by the position error at the end of the run (hand measured to be 10 cm) which lies within the 3σ bounds.

V. Conclusions and Future Work

This paper presents a novel approach to indoor localization for the visually impaired. In our proposed estimation scheme, information from a pair of cane-mounted sensors, and a foot-mounted pedometer is fused in a two-layer pose estimator. The first layer utilizes inertial measurements from a 3-axis gyroscope and relative orientation measurements from laser scanner data to accurately estimate the attitude of the white cane. The second layer estimates the position of the person holding the cane, by processing linear velocity measurements from the pedometer, a filtered version of the cane's yaw estimate, and corner features extracted from the laser scanner data.

A desirable feature of our system is that it is lightweight and unobtrusive, and since it is cane mounted, the person maintains the ability to physically touch the environment. We demonstrate the validity of our localization aid with experimental results showing that we can provide accurate pose estimates of the person even in dynamic environments. Our system requires no building instrumentation. Instead, we utilize a corner map which can be obtained from the building blueprints or generated by a robot exploring the environment beforehand. Furthermore, we analyze the observability properties of the system and provide a sufficient condition which ensures that we can accurately estimate the 3D orientation of the white cane.

In our view, the described white-cane based localization system has the potential to spearhead the development of indoor navigation and guidance aids for the visually impaired. As part of our future work, we plan to extend the proposed system and algorithm to support navigation in unknown buildings while concurrently constructing a map of the area. We also envision an Internet-based map distribution system in which different visually impaired users can collaborate and refine maps which their white canes have constructed, adding contextual and semantic information. Our future work also includes the design of path planning routines to generate routes between indoor locations, as well as a haptic feedback system to give the person simple directions. A software implementation on a small-scale computing device such as a personal digital assistant, or an embedded computer is also within our short term goals.

Appendix A

In this appendix, we provide a brief overview of the weighted line fitting method that we employ for extracting lines from the raw laser scan data (Pfister et al. 2003). This method exploits the uncertainty in each laser point to segment the laser scan into multiple lines and compute an estimate of the line parameters for each line in the scan. We also discuss issues of robustness and reliability of the line extraction algorithm for detecting the structural planes as well as corner features in the environment.

When a new laser scan is recorded, the raw data is partitioned into the even-indexed and odd-indexed points, which will be utilized in the first and second layers of the estimator, respectively. We process the two halves of the data independently to ensure that the two layers of the filter remain statistically uncorrelated. If we were to reuse the same laser data in both layers, then their estimates would be correlated. Although there exist techniques for dealing with such correlations (Mourikis et al. 2007), for the sake of simplicity we have opted to use only half of the points in each layer.

A. Scan Segmentation

The first step of the weighted line fitting algorithm is to segment the data into groups of points which belong to the same lines. To accomplish this task, we employ the standard Hough Transform (HT), which is a voting-based scheme that identifies groups of points that are roughly collinear. A 2D grid is constructed in which every grid cell corresponds to the polar coordinates of a single line. Each laser data point votes for the line that it most closely matches, and the grid cells with the most votes are taken as possible lines in the laser scan data (Duda and Hart 1972). We note that the results of the HT need not be perfect since we will refine each line estimate, and subsequently merge similar lines segments based on a probabilistic matching procedure.

B. Line Fitting

Given a set of N points that are roughly collinear (as determined by the HT), the weighted line fitting problem seeks to determine the optimal line parameters (ρ^*, ϕ^*) that maximize the probability that each point lies on the line. The measured laser points are described by their polar coordinates $(d_{m,j}, \theta_{m,j})$, $j = 1, \ldots, N$, which correspond to noise-corrupted versions of the true points that lie on the line of intersection between the laser scanning plane and a structural plane, i.e.,

$$d_{j,m} = d_j + \eta_d \tag{37}$$

$$\theta_{j,m} = \theta_j + \eta_\theta \tag{38}$$

where η_d and η_θ are zero-mean white Gaussian random variables with variances σ_d^2 and σ_θ^2, respectively. We define the distance from a measured point to the line as

$$\epsilon_j = d_{m,j} \cos(\phi - \theta_{m,j}) - \rho. \tag{39}$$

From this, we formulate a Maximum Likelihood Estimator (MLE) to determine (ρ^*, ϕ^*), i.e.,

$$\begin{aligned}(\rho^*, \phi^*) &= \underset{\rho,\phi}{\operatorname{argmax}} \mathcal{L}(\{\epsilon_j\}|\rho, \phi) \\ &= \underset{\rho,\phi}{\operatorname{argmax}} \mathcal{L}(\epsilon_1|\rho, \phi)\, \mathcal{L}(\epsilon_2|\rho, \phi) \cdots \mathcal{L}(\epsilon_N|\rho, \phi)\end{aligned} \tag{40}$$

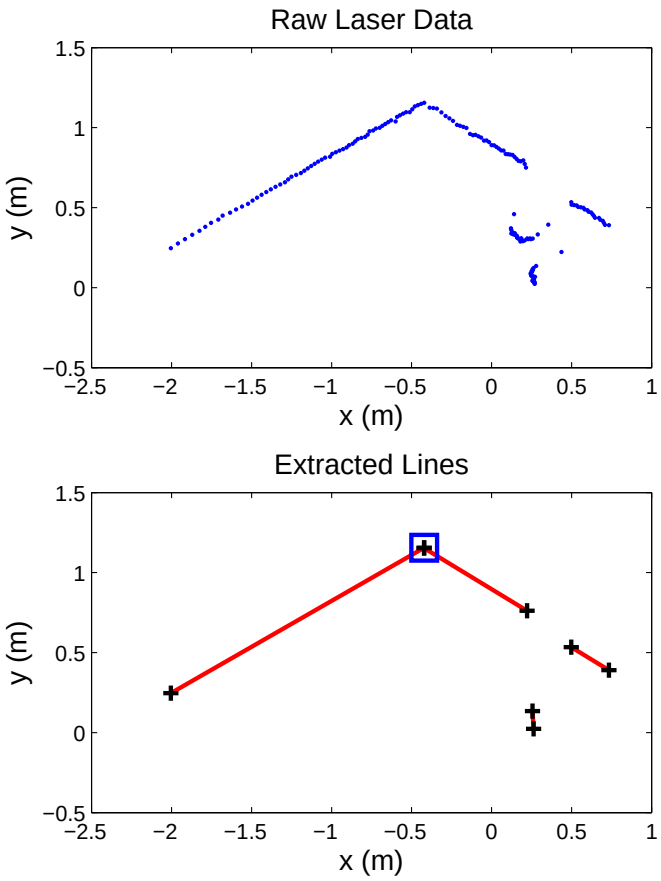

Fig. 13. (top) Raw laser scan in which the laser scanner was pointed into a corner. The laser scan plane intersects with a person's legs, as well as with the right and left walls. Since the laser scanner is cane-mounted and inclined with respect to the world, the 90 deg corner appears obtuse in the data. (bottom) The lines extracted from the laser data (red line segments between black 'plus' sign endpoints), the two walls are correctly extracted, and one of the person's legs is also extracted, but discarded later on since it is very short. The corner at the intersection of the left and right walls is also identified (boxed in blue).

where $\mathcal{L}(\{\epsilon_j\}|\rho,\phi)$ is the joint likelihood of ϵ_j, $j = 1,\ldots,N$ given the line parameters. Based on the assumption of independently distributed Gaussian noise [see (37) and (38)], we factorize the joint likelihood into the product of the likelihoods for each ϵ_j conditioned on the line parameters. Computing the negative log-likelihood of (40), we obtain an equivalent minimization problem that is solved using nonlinear weighted least squares. We refer the interested reader to (Pfister et al. 2003) for more details.

C. Line Merging

From the previous section, we obtain a set of lines, where for each line, we computed the MLE estimate of the line parameters based on the group of points assigned to that line. Due to the discretization in line space during the HT step, it may happen that laser data points corresponding to the same physical line were split into two or more groups. We perform a line merging procedure in which we test if two adjacent estimated lines correspond to the same physical line with a probabilistic test based on the uncertainty in both line-parameter estimates. If two estimated lines are found to correspond to the same physical line, then a new line is computed by merging the line estimates. The closed-form solution to compute the line parameters and covariance for the merged line are given in (Pfister et al. 2003).

D. Robustness and Reliability

There are several key observations that should be made about the robustness and reliability of the line fitting procedure. First, we cannot detect every line in the laser data. In fact, the HT-based methods have been empirically evaluated and found to have a true positive rate around 80% (Nguyen et al. 2005). However, for the lines that are detected, we obtain the optimal estimate of the line parameters (in the MLE sense). Second, we will have false positives, which are detected lines that do not correspond to lines of intersection between the structural planes and the laser scan plane, but instead correspond to nearby furniture or obstacles (see Fig. 13). However, detecting incorrect lines is typically not an issue, since these lines are discarded during the data association step (i.e., they will not match one of the structural planes of the building). To aid this, we set thresholds on the minimum number of points per line and the minimum line length that we will accept. Third, the robustness of the line extraction and data association step enable reliable corner extraction. Corners are only declared if two line segments in the scan have been gated, and their end points lie within a threshold distance of each other. Furthermore, after each corner is detected, we perform a probabilistic data association test to verify that it matches one of the map corners before it is incorporated into the filter.

Fig. 13 illustrates typical results from the laser scan fitting procedure. Even in the presence of an obstacle (i.e., person's legs), the left and right walls are correctly extracted. Although one of the legs is identified as a short line segment, this measurement is rejected by the filter's data association step. The corner at the intersection of the left and right walls is also identified in this scan.

REFERENCES

W. Balachandran, F. Cecelja, and P. Ptasinski. A GPS based navigation aid for the blind. In *Proc. of the Int. Conf. on Applied Electromagnetics and Communications*, pages 34–36, Dubrovnik, Croatia, Oct. 1–3, 2003.

D. A. Bell, J. Borenstein, S. P. Levine, Y. Koren, and L. Jaros. An assistive navigation system for wheelchairs based on mobile robot obstacle avoidance. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 2018–2022, San Diego, CA, May 8–13, 1994.

W. G. Breckenridge. Quaternions proposed standard conventions. Interoffice Memorandum IOM 343-79-1199, Jet Propulsion Laboratory, Pasadena, CA, 1999.

F. Cavallo, A. 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.

S. Ceranka and M. Niedzwiecki. Application of particle filtering in navigation system for blind. In *Proc. of the Int. Sym. on Signal Processing and its Applications*, pages 495–498, Paris, France, July 1–4, 2003.

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.

F. Dellaert, D. Fox, W. Burgard, and S. Thrun. Monte carlo localization for mobile robots. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 1322–1328, Detroit, MI, May 10–15, 1999.

R. O. Duda and P. E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, Jan. 1972.

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.

N. A. Giudice and G. E. Legge. Blind navigation and the role of technology. In A. Helal, M. Mokhtari, and B. Abdulrazak, editors, *The Engineering Handbook of Smart Technology for Aging, Disability and Independence*, Computer Engineering Series, chapter 25. John Wiley & Sons, 2008.

J. A. Hesch and S. I. Roumeliotis. An indoor navigation aid for the visually impaired. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 3545–3551, Rome, Italy, Apr. 10–14, 2007.

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. BrailleNote global positioning system, 2009. URL `http://www.humanware.com/`.

IEEE Std. 647-2006. IEEE standard specification format guide and test procedure for single-axis laser gyros. IEEE Aerospace and Electronic Systems Society, New York, ISBN: 0-7381-4950-0, Sept. 2006.

A. Kelly. Linearized error propagation in odometry. *Int. Journal of Robotics Research*, 23(2):179–218, Feb. 2004.

R. L. Knoblauch, M. T. Pietrucha, and M. Nitzburg. Field studies of pedestrian walking speeds and start-up time. *Journal of the Transportation Research Board*, 1538:27–38, 1996.

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, 2004a.

V. Kulyukin, C. Gharpure, P. Sute, N. De Graw, and J. Nicholson. A robotic wayfinding system for the visually impaired. In *Proc. of the Innovative Applications of Artificial Intelligence Conf.*, pages 864–869, San Jose, CA, July 27–29, 2004b.

J. J. Leonard and H. F. Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *IEEE Trans. on Robotics and Automation*, 7(3):376–382, June 1991.

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

P. S. Maybeck. *Stochastic models, estimation, and control*. Academic Press, 1979.

P. B. L. Meijer. An experimental system for auditory image representations. *IEEE Trans. on Biomedical Engineering*, 39(2):112–121, Feb. 1992.

M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit. FastSLAM: a factored solution to the simultaneous localization and mapping problem. In *Proc. of the AAAI National Conf. on Artificial Intelligence*, pages 593–598, Edmonton, Canada, July 28 – Aug. 1, 2002.

A. I. Mourikis and S. I. Roumeliotis. Performance bounds for cooperative simultaneous localization and mapping (C-SLAM). In *Proc. of Robotics: Science and Systems*, pages 73–80, Cambridge, MA, June 8–11, 2005.

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.

P. Moutarlier and R. Chatila. Stochastic multisensory data fusion for mobile robot location and environment modeling. In *Proc. of the Int. Sym. on Robotics Research*, Tokyo, Japan, Aug. 1989.

V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart. A comparison of line extraction algorithms using 2D laser rangefinder for indoor mobile robotics. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1929–1934, Edmonton, Canada, Aug. 2–6, 2005.

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.

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 Group. Mobile geo navigation software, 2009. URL `http://www.senderogroup.com/`.

I. Shim and J. Yoon. Shared behavior control of a robotic cane based on interactive technology. In *Proc. of the IEEE Conf. of the Industrial Electronics Society*, pages 2623–2628, Sevilla, Spain, Nov. 5–8, 2002.

S. Shoval, J. Borenstein, and Y. Koren. Auditory guidance with the navbelt - a computerized travel aid for the blind. *IEEE Trans. on Systems, Man, and Cybernetics*, 28(3):459–467, Aug. 1998.

R. Smith, M. Self, and P. Cheeseman. *Autonomous Robot Vehicles*, chapter Estimating Uncertain Spatial Relationships

in Robotics, pages 167–193. Springer-Verlag, 1990.

S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT Press, Sept. 2005.

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*, page 30, San Diego, CA, June 20–25, 2005.

N. Trawny and S. I. Roumeliotis. Indirect Kalman filter for 3D attitude estimation. Technical Report 2005-002, University of Minnesota, Dept. of Comp. Sci. & Eng., MARS Lab, 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.

G. Whitney. New technology assistive devices to assist blind and partially sighted people to travel within and use the built environment. In *Proc. of the Int. Conf. on Road Transport Information and Control*, pages 55–58, London, UK, Apr. 21–23, 1998.

S. B. Williams, G. Dissanayake, and H. Durrant-Whyte. An efficient approach to the simultaneous localization and mapping problem. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 406–411, Washington, DC, May 11–15, 2002.

F. Wong, R. Nagarajan, and S. Yaacob. Application of stereo vision in a navigation aid for blind people. In *Proc. of the Joint Conf. of the Int. Conf. on Information, Communications and Signal Processing and the Pacific-Rim Conf. on Multimedia*, pages 734–737, Singapore, Dec. 15–18, 2003.

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.

X. Yun, E. R. Bachmann, H. Moore IV, and J. Calusdian. Self-contained position tracking of human movement using small inertial/magnetic sensor modules. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, pages 2526–2533, Rome, Italy, Apr. 10–14, 2007.

J. S. Zelek, S. Bromley, D. Asmar, and D. Thompson. A haptic glove as a tactile-vision sensory substitution for wayfinding. *Journal of Visual Impairment and Blindness*, 97(10):621–632, Oct. 2003.