2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

An Indoor Localization Aid for the Visually Impaired

Joel A. Hesch and Stergios I. Roumeliotis

Abstract— **This paper presents an indoor human localization system for the visually impaired. A prototype portable device has been implemented, consisting of a pedometer and a standard white cane, on which a laser range finder and a 3-axis gyroscope have been mounted. A novel pose estimation algorithm has been developed for robustly estimating the heading and position of a person navigating in a known building. The basis of our estimation scheme is a two-layered Extended Kalman Filter (EKF) for attitude and position estimation. The first layer maintains an attitude estimate of the white cane, which is subsequently provided to the second layer where a position estimate of the user is generated. Experimental results are presented that demonstrate the reliability of the proposed method for accurate, real-time human localization.**

I. Introduction

Mobility is an essential capability for any person who wishes to have 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 these tasks may be exceedingly difficult to accomplish, and there are high risks associated with failure in any of these. Seeing-eye dogs and white canes are widely used for the purpose of guidance and environment sensing. The former, however, has costly and often prohibitive training requirements, while the latter can only provide cues about ones immediate surroundings. Human performance on information-dependant tasks, can be improved by sensing which provides information (e.g., position, orientation, or local geometry) and environmental cues via the use of appropriate sensors and sensor fusion algorithms. This paper presents a novel indoor localization method for the visually impaired which has the potential for prodigious humanitarian impact. With the use of this localization aid, guidance and navigation algorithms can be implemented which will greatly increase the safety and overall mobility of its user.

When designing a suitable sensor package for use in a human localization application, the 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 [1], [2]. 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 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 (e.g., sitting in a chair), and may prevent certain articles of clothing from being comfortably worn (e.g., a jacket). In contrast to this approach, we propose using a sensor package *mounted on a white cane* (cf. Fig. 3). The main advantages of utilizing a white cane as a sensor platform are: (i) the sensor package is *unobtrusive* to the user, (ii) there is *no* need to calibrate the system for specific body types, and (iii) the user maintains the ability to *physically touch* the environment with the white cane.

After considering the proposed sensor placement, one can appreciate the stark difference between indoor human localization, and traditional formulations of mobile (wheeled) robot localization. When constructing an estimator for the pose of a mobile robot, accurate linear and rotational velocity measurements are available from its wheel encoders. In the case of a blind person carrying a white cane, the sensors providing these measurements are not rigidly connected, which makes the task of combining information from them significantly more challenging. Specifically, a pair of sensors (a laser scanner and a 3-axis gyroscope) mounted under the white cane handle provide *attitude* information about the *cane*, and a lightweight, foot-mounted pedometer measures the *user's* walking *speed*. Ideally, information from these three sources should be fused in a single pose estimator. All the sensors, however, move in 3-D and the coordinate transformation from the pedometer to the laser/gyro is *unknown* and *time varying*.

In order to address this problem, we have designed and implemented a two-layered (2.5-D) estimator. In the first stage, *rotational velocity* measurements from the 3-axis gyroscope are combined with *relative attitude* measurements inferred from the laser scan data to estimate the 3-D *attitude* of the cane. The second stage incorporates *corner features* extracted from the laser data, *linear velocity* measurements from the pedometer, and a filtered version of the cane's *yaw* to compute 2-D position estimates of the user.[1] By exploiting *a priori* information about the location of environmental features (corners), and considering that many of the primary structural planes (floor, ceiling, walls) of a building lie perpendicular to each other, the described method generates a reliable localization estimate of a person traveling indoors.

Section II of this paper reviews the relevant literature on obstacle avoidance, navigation, and localization systems for the visually impaired. The problem of estimating the 3-D attitude of the white cane is discussed in Section III-A. Section III-B details the use of a low-pass filter to extract the *heading* of the user from the attitude estimate of the cane.

This work was supported by the University of Minnesota (GiA program) and the National Science Foundation (EIA-0324864, IIS-0643680).

The authors are with the Dept. of Computer Science & Engineering, University of Minnesota, Minneapolis, MN 55455. Emails: {joel|stergios}@cs.umn.edu

[1] It is important to note that while the laser data are utilized in both stages of the filter, statistical correlations in the estimates are avoided by using the even-indexed data points in the first stage, and the odd-indexed data points in the second stage.

The 2-D position filter for estimating the user's location is presented in Section III-C. A description of the hardware utilized is given in Section IV-A. Experimental results of the method presented here are provided in Section IV-B. Finally, 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 [3]. These employ sensors for *obstacle avoidance* such as laser pointers [4], and sonars on a wheelchair [5], on a robot connected at the tip of a white cane [6], [7], or as part of a travel aid [1], [2]. Cameras have also been suggested [8], [9] for *object description* (in terms of color and size) in addition to obstacle detection. While these devices augment the cognitive abilities of a blind person and reduce the probability of an accident due to an undetected obstacle, they cannot be explicitly used as a *wayfinding* aid without the development of appropriate algorithms for localization and mapping.

Significant research work (e.g., [10], [11]) has concentrated on mobile robot navigation. However, there are only few attempts to apply this knowledge to assist visually impaired people in their everyday navigation tasks. Instead most relevant efforts have focused primarily on GPS-based *outdoor navigation* (e.g., [12], [13], [14], [2]) which cannot be used inside a building. An approach to *indoor wayfinding* for the visually impaired is presented in [15], [16], [17]. In this case, an autonomous robot attached at the end of a leash, as a substitute for a guide dog, 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 crowded rooms. Additionally, the weight and volume of the robot, negatively affects its portability by a commuter. Furthermore, it requires instrumentation of buildings with RFIDs which is costly and time consuming; this is also the case for similar ultrasound [2] and Infra Red (IR) [18] based systems. In contrast, we are interested in designing a white cane-mounted sensor system to aid visually impaired people for indoor navigation. This is more challenging due to the variations in body geometry (e.g., height and stride) and motion patterns across different people. Additional difficulties arise when dealing with cane-based sensor systems due to the unknown and time-varying coordinate transformation between the sensors and the user. However, the white cane is an ideal platform for indoor human localization for several reasons: (i) it is a trusted tool, already in use by the target demographic, (ii) it is lightweight, portable, and unobtrusive to the user, and (iii) a cane-based localization system requires no building instrumentation.

III. Method Description

The algorithm described in this work consists of three main components. First, the attitude of the white cane is estimated using a 3-axis gyroscope, and laser-scan measurements of structural planes in the building. Second, the heading direction of the person is extracted from the yaw component of the white cane's attitude estimate with the use of a low-pass filter. The purpose of this step is to provide a heading measurement to the second stage of the filter. Lastly, the position of the person is estimated using the heading estimates from the low-pass filter, the linear velocity measurements from the pedometer, and the relative coordinates of known corner features detected by the laser scanner.

A. Attitude Estimation of the White Cane

In this work, attitude is represented using the quaternion of rotation:

$$\mathbf{q} = {}^{S}_{G}\mathbf{q} = \begin{bmatrix} \widehat{\mathbf{k}} \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}^{\mathrm{T}} \tag{1}$$

where $\{S\}$ and $\{G\}$ denote the gyroscope and global frames of reference, $\widehat{\mathbf{k}}$ is the axis of rotation, and θ signifies its magnitude. 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 ${}^{L_1}_{L_3}q = {}^{L_1}_{L_2}q \otimes {}^{L_2}_{L_3}q$, which is the attitude rotation between successive frames [19].

Attitude estimation is accomplished through the use of an EKF which fuses measurements from proprioceptive and exteroceptive sensing devices. Rotational velocity measurements from a 3-axis gyroscope are integrated to propagate the attitude estimate, and straight lines extracted from the laser-scan data are used to update the computed estimate.

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

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{q} \\ \mathbf{b} \end{bmatrix} \quad , \quad \tilde{\mathbf{x}}_k = \begin{bmatrix} \delta\boldsymbol{\theta} \\ \widetilde{\mathbf{b}} \end{bmatrix}$$

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

$$\delta\mathbf{q} = \mathbf{q} \otimes \widehat{\mathbf{q}}^{-1} \simeq \begin{bmatrix} \frac{1}{2}\delta\boldsymbol{\theta}^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}} \tag{2}$$

The error quaternion $\delta\mathbf{q}$ denotes a small rotational error between the true, $\mathbf{q}$, and the estimated, $\widehat{\mathbf{q}}$, attitude of the cane.

a) Continuous-time model: The state model for the quaternion representation of attitude is governed by the quaternion time derivative (3), which is computed in terms of the instantaneous rotational velocity $\boldsymbol{\omega}$.

$$\dot{\mathbf{q}}(t) = \frac{1}{2}\boldsymbol{\Omega}(\boldsymbol{\omega})\,\mathbf{q}(t) \tag{3}$$

where

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

The gyroscope measures the rotational velocity of the cane expressed with respect to the local (cane-affixed) frame of reference. These measurements are corrupted by sensor bias $\mathbf{b}$, as well as measurement noise $\mathbf{n}_r$:

$$\boldsymbol{\omega}_m = \boldsymbol{\omega} + \mathbf{b} + \mathbf{n}_r \tag{4}$$

where the turn-rate noise $\mathbf{n}_r$ is distributed as zero-mean white Gaussian process with covariance $\sigma_r^2\mathbf{I}_{3\times3}$. The sensor bias $\mathbf{b}$ is modeled as a random walk with

$$\dot{\mathbf{b}} = \mathbf{n}_w \tag{5}$$

where $\mathbf{n}_w$ is also distributed as zero-mean white Gaussian noise process with covariance $\sigma_w^2\mathbf{I}_{3\times3}$.

The continuous-time error-state propagation is described by the following equation [20]:

$$\begin{bmatrix}\dot{\delta\boldsymbol{\theta}}\\ \dot{\tilde{\mathbf{b}}}\end{bmatrix} = \begin{bmatrix}-\lfloor\widehat{\boldsymbol{\omega}}\times\rfloor & -\mathbf{I}_{3\times3}\\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3}\end{bmatrix}\begin{bmatrix}\delta\boldsymbol{\theta}\\ \tilde{\mathbf{b}}\end{bmatrix} + \begin{bmatrix}-\mathbf{I}_{3\times3} & \mathbf{0}_{3\times3}\\ \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3}\end{bmatrix}\begin{bmatrix}\mathbf{n}_r\\ \mathbf{n}_w\end{bmatrix}$$
$$\dot{\tilde{\mathbf{x}}} = \mathbf{F}_c\cdot\tilde{\mathbf{x}} + \mathbf{G}_c\cdot\mathbf{n} \tag{6}$$

where $\mathbf{F}_c$ is the continuous-time system matrix, $\mathbf{G}_c$ is the input noise matrix, and $\mathbf{0}_{m\times n}$ and $\mathbf{I}_{m\times n}$ are the $m\times n$ zero, and identity matrices, respectively.

b) Discrete-time implementation: During each propagation step, the bias estimate is considered constant (cf. (5))

$$\widehat{\mathbf{b}}_{k+1|k} = \widehat{\mathbf{b}}_{k|k} \tag{7}$$

and the quaternion estimate is propagated by integrating (3):

$$\widehat{\mathbf{q}}_{k+1|k} = \begin{bmatrix}\frac{\widehat{\boldsymbol{\omega}}_{k|k}}{||\widehat{\boldsymbol{\omega}}_{k|k}||}\sin\left(\frac{||\widehat{\boldsymbol{\omega}}_{k|k}||}{2}\delta t\right)\\ \cos\left(\frac{||\widehat{\boldsymbol{\omega}}_{k|k}||}{2}\delta t\right)\end{bmatrix}\otimes\widehat{\mathbf{q}}_{k|k} \tag{8}$$

where $||\widehat{\boldsymbol{\omega}}_{k|k}|| = \sqrt{\widehat{\boldsymbol{\omega}}_{k|k}^{\mathrm{T}}\widehat{\boldsymbol{\omega}}_{k|k}}$ and $\widehat{\boldsymbol{\omega}}_{k|k} = \boldsymbol{\omega}_m(t_k) - \widehat{\mathbf{b}}_{k|k}$.

Finally, the error-state covariance matrix is propagated as:

$$\mathbf{P}_{k+1|k} = \boldsymbol{\Phi}_k\mathbf{P}_{k|k}\boldsymbol{\Phi}_k^{\mathrm{T}} + \mathbf{Q}_{d_k} \tag{9}$$

where

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

and

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

with $\mathbf{Q}_c = \begin{bmatrix}\sigma_r^2\mathbf{I}_{3\times3} & \mathbf{0}_{3\times3}\\ \mathbf{0}_{3\times3} & \sigma_w^2\mathbf{I}_{3\times3}\end{bmatrix}$.

2) Attitude Update: Even-indexed laser scan data points from a laser range finder are employed to measure the relative orientation between the sensor frame of reference $\{S\}$ and the global frame $\{G\}$. Specifically, the laser sensing plane intersects the planar surfaces inside a building (e.g., walls, floor, ceiling) along straight lines which can be reliably detected and extracted from the laser data. The *direction* of each of the extracted lines is processed as a measurement for updating the attitude estimates.

Inside a building, frame $\{G\}$ can be assigned such that its principal axes $\{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ are perpendicular to the prominent structural planes of the building (i.e., $\mathbf{e}_1 \perp Wall_x$, $\mathbf{e}_2 \perp Wall_y$, and $\mathbf{e}_3 \perp Floor, Ceiling$). Measurements to these planes are denoted by their corresponding unit-vectors (e.g., an x-measurement is a measured line which is perpendicular to $\mathbf{e}_1$). Let $\mathbf{e}_i \in \{\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3\}$ be one of the three unit vectors of frame $\{G\}$, and let ${}^G\boldsymbol{\ell}$ denote the (unit vector) direction of the line of intersection between the laser-scan plane and the measured plane expressed with respect to $\{G\}$. From the geometric constraint, their inner product should be zero

$$\mathbf{e}_i^{\mathrm{T}\,G}\boldsymbol{\ell} = 0,\ \ i = 1,2,\text{ or }3 \tag{10}$$

The inferred measurement equation is derived by rewriting this constraint using the transformation relation ${}^G\boldsymbol{\ell} = C^{\mathrm{T}}(\mathbf{q})\,{}^S\boldsymbol{\ell}$:

$$z = \mathbf{e}_i^{\mathrm{T}}C^{\mathrm{T}}(\mathbf{q})\,{}^S\boldsymbol{\ell} = 0 \tag{11}$$

where the rotation matrix $C^{\mathrm{T}}(\mathbf{q})$ projects vectors expressed with respect to frame $\{S\}$ to frame $\{G\}$. Note that since ${}^S\boldsymbol{\ell}$ is the unit vector direction of a line on the x-y plane of the laser sensor frame, it can be written as ${}^S\boldsymbol{\ell} = \begin{bmatrix}\sin\phi & -\cos\phi & 0\end{bmatrix}^{\mathrm{T}}$ where ϕ is the complimentary angle to the line direction.

The estimated measurement equation is

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

where ${}^S\boldsymbol{\ell}_m = {}^S\boldsymbol{\ell} - {}^S\tilde{\boldsymbol{\ell}}$ denotes the measured line direction and ${}^S\tilde{\boldsymbol{\ell}} = \lfloor\mathbf{e_3}\times\rfloor{}^S\boldsymbol{\ell}_m\tilde{\phi}$ is the error in this measurement. Note that the line direction error $\tilde{\phi}\sim\mathcal{N}(0,\sigma_\phi^2)$ accounts for measurement noise as well as inaccuracies in line fitting.

Employing (11) and (12), the measurement error is:

$$\tilde{z} \simeq \begin{bmatrix}-\mathbf{e}_i^{\mathrm{T}}C^{\mathrm{T}}(\widehat{\mathbf{q}})\lfloor{}^S\boldsymbol{\ell}_m\times\rfloor & \mathbf{0}_{1\times3}\end{bmatrix}\begin{bmatrix}\delta\boldsymbol{\theta}\\ \tilde{\mathbf{b}}\end{bmatrix} + n$$
$$= \mathbf{h}^{\mathrm{T}}\tilde{\mathbf{x}} + n \tag{13}$$

where $n = \mathbf{e}_i^{\mathrm{T}}C^{\mathrm{T}}(\widehat{\mathbf{q}})\lfloor\mathbf{e_3}\times\rfloor{}^S\boldsymbol{\ell}_m\tilde{\phi}$.

Updating the state estimates requires to compute the residual

$$r = z - \mathbf{e}_i^{\mathrm{T}}C^{\mathrm{T}}(\widehat{\mathbf{q}})\,{}^S\boldsymbol{\ell}_m = -\mathbf{e}_i^{\mathrm{T}}C^{\mathrm{T}}(\widehat{\mathbf{q}})\,{}^S\boldsymbol{\ell}_m \tag{14}$$

and the Kalman gain:

$$\mathbf{k} = \mathbf{P}_{k+1|k}\mathbf{h}\left(\mathbf{h}^{\mathrm{T}}\mathbf{P}_{k+1|k}\mathbf{h} + R\right)^{-1} \tag{15}$$

where $R = E\{n^2\} = \left(\mathbf{e}_i^{\mathrm{T}}C^{\mathrm{T}}(\widehat{\mathbf{q}})\lfloor\mathbf{e_3}\times\rfloor{}^S\boldsymbol{\ell}_m\right)^2\sigma_\phi^2$.

The error-state estimate is given by the following expression

$$\widehat{\mathbf{x}}(+) = \begin{bmatrix}\delta\widehat{\boldsymbol{\theta}}(+)\\ \widehat{\mathbf{b}}(+)\end{bmatrix} = \mathbf{k}r \tag{16}$$

From $\delta\widehat{\boldsymbol{\theta}}(+)$ the error quaternion estimate is computed as

$$\delta\widehat{\mathbf{q}} = \frac{1}{\sqrt{1+\frac{1}{4}\delta\widehat{\boldsymbol{\theta}}(+)^{\mathrm{T}}\delta\widehat{\boldsymbol{\theta}}(+)}}\cdot\begin{bmatrix}\frac{1}{2}\delta\widehat{\boldsymbol{\theta}}(+)\\ 1\end{bmatrix} \tag{17}$$

The quaternion and the bias estimates are updated as

$$\widehat{\mathbf{q}}_{k+1|k+1} = \delta\widehat{\mathbf{q}}\otimes\widehat{\mathbf{q}}_{k+1|k} \tag{18}$$

$$\widehat{\mathbf{b}}_{k+1|k+1} = \widehat{\mathbf{b}}_{k+1|k} + \widehat{\mathbf{b}}(+) \tag{19}$$

The last step in this process is the covariance update:

$$\mathbf{P}_{k+1|k+1} = (\mathbf{I} - \mathbf{k}\mathbf{h}^{\mathrm{T}})\mathbf{P}_{k+1|k}(\mathbf{I} - \mathbf{k}\mathbf{h}^{\mathrm{T}})^{\mathrm{T}} + \mathbf{k}R\mathbf{k}^{\mathrm{T}} \tag{20}$$

3) Attitude Observability: Due to physical limitations of *planar* laser scanning, it is impossible to resolve all three degrees of rotational freedom from a *stationary* vantage point. The attitude of the cane, however, is *stochastically observable* when the cane is in *motion* [21]. In what follows, we prove that the observability requirements are satisfied when the laser sensor detects all three main directions in space, at different time steps, over a given time interval.

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

$$\mathbf{h}_i^T(t_k) = -\mathbf{e}_i^{\mathrm{T}} C^{\mathrm{T}}\left(\mathbf{q}(t_k)\right) \lfloor {}^S\boldsymbol{\ell}_i(t_k)\times \rfloor \tag{21}$$

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

$$\mathbf{M} = \sum_{k=1}^{N}\sum_{i} \boldsymbol{\Phi}^{\mathrm{T}}(t_k,0)\mathbf{h}_i(t_k)\mathbf{h}_i^{\mathrm{T}}(t_k)\boldsymbol{\Phi}(t_k,0) \tag{22}$$

is of full rank for some finite N [21].

Noting that in this case the state transition matrix $\boldsymbol{\Phi}(t_k,0) = \mathbf{C}(\mathbf{q}(t_k))$, the observability Gramian can be written as the sum of the following vector outer products:

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

with ${}^G\mathbf{h}_i(t_k) = {}^G\boldsymbol{\ell}_i(t_k) \times \mathbf{e}_i$. If over a period of time the sensor observes surfaces with normals $\mathbf{e}_i$ that span $\mathbb{R}^3$, and recalling that ${}^G\boldsymbol{\ell}_i^{\mathrm{T}}\mathbf{e}_i = 0$ (cf. (10)), then the vectors ${}^G\mathbf{h}_i$ also span the 3D-space, ensuring $\mathbf{M}$ to be of full rank.

Fig. 1 depicts the trace of the attitude covariance matrix with respect to time. During this experiment, the white cane was initially stationary for 10sec for the purpose of gyroscope bias initialization. As evident from the experimental results, the trace of the attitude covariance becomes bounded (as expected since the system is observable) once the cane is in motion.

B. Heading Estimation of the Person

During regular operation, the yaw angle of the cane will be an asymmetric cyclostationary random process in which the amplitude, phase, frequency, and degree of asymmetry may all change (cf. Fig. 2). These characteristics result from swinging the white cane which helps the person identify a clear walking path. Extracting the heading direction of the person from the cane's yaw can be accomplished through the employment of traditional signal processing techniques. However, due to the constraint that the filter must propagate in real time, a suboptimal method has been employed.

The heading of a person can be well approximated as the mean value of the cane yaw over a period of swinging. Thus, a proper solution would be a combination of frequency estimation, and low pass filtering to remove the high frequency

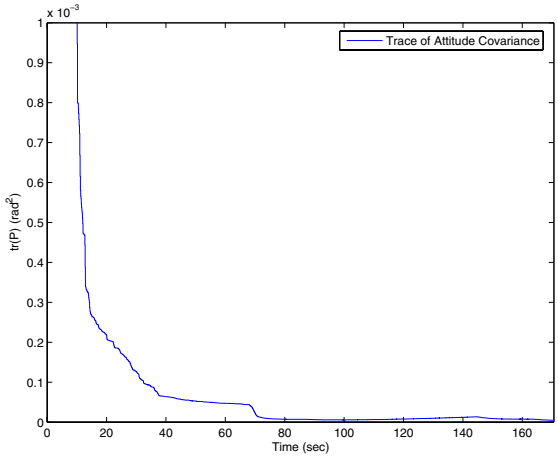

Fig. 1. The trace of the attitude covariance demonstrates that while the cane is stationary (first 10sec during bias initialization) its attitude is initially unobservable, however, it becomes observable when the cane is in motion.

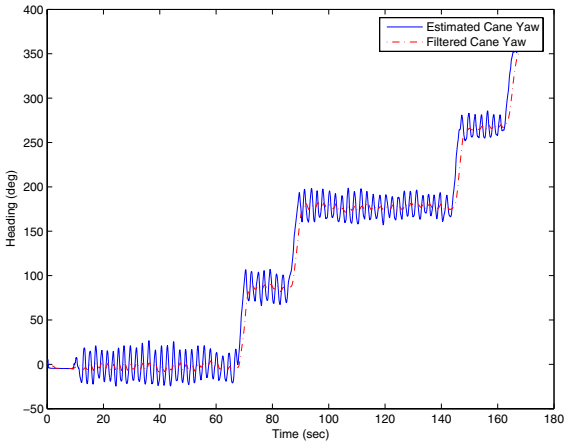

Fig. 2. The yaw component of the cane's attitude estimate plotted before and after low-pass filtering.

component of the yaw. This motivates the use of a Finite Impulse Response (FIR) filter. The attitude is propagated at 100Hz, however, due to the nature of the motion of the cane the high frequency component of the yaw does not fall below 0.5Hz. In order to reduce the number of taps needed by the 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 0.02. Fig. 2 depicts the yaw component of the cane's attitude estimate along with the filtered version. A byproduct of filtering the yaw signal is that a delay is introduced. Although the filter has only 7 coefficient-delay pairs, there is a delay of 1.2sec because the down-sampled yaw signal has a sampling frequency of 2.5Hz. As a result, an equivalent time-lag exists in the position estimate. However, due to the relatively slow walking rate of a person carrying a white cane, this delay is not prohibitive for real-time operation.

C. Position Estimation of the Person

Estimating the position of a person within a building can be treated as a 2-D position estimation problem in

which each floor of the building is a separate environment containing landmarks, in this case corner features, whose position is known. 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 2-D 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* [22]. The state propagation equations are:

$$\begin{aligned} x_{k+1} &= x_k + V\delta t c\psi \quad , \quad y_{k+1} = y_k + V\delta t s\psi \\ \widehat{x}_{k+1} &= \widehat{x}_k + V_m\delta t c\psi_m \quad , \quad \widehat{y}_{k+1} = \widehat{y}_k + V_m\delta t s\psi_m \end{aligned}$$

where $(x,\ y)$ and ψ are the position and heading of the person, and V is the average velocity during the time interval δt. In the above equations " $\widehat{\ }$ " denotes estimates while the subscript m refers to measured quantities, i.e.,

$$V_m = V + w_v \qquad , \qquad \psi_m = \psi + w_\psi$$

where the velocity, w_v, and heading, w_ψ, errors are zero-mean white Gaussian processes with variances σ_v^2 and σ_ψ^2, respectively. The error model based on these relations is:

$$\begin{aligned} \begin{bmatrix} \tilde{x}_{k+1} \\ \tilde{y}_{k+1} \end{bmatrix} &= \begin{bmatrix} \tilde{x}_k \\ \tilde{y}_k \end{bmatrix} + \begin{bmatrix} -\delta t c\psi_m & V_m\delta t s\psi_m \\ -\delta t s\psi_m & -V_m\delta t c\psi_m \end{bmatrix} \begin{bmatrix} w_v \\ w_\psi \end{bmatrix} \\ \tilde{\mathbf{x}}_{k+1} &= \tilde{\mathbf{x}}_k + \mathbf{\Gamma}\mathbf{w} \end{aligned} \tag{23}$$

and the covariance propagation is computed as

$$\mathbf{P}_{k+1|k} = \mathbf{P}_{k|k} + \mathbf{\Gamma}\mathbf{Q}\mathbf{\Gamma}^{\mathrm{T}}, \ \text{where } \mathbf{Q} = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_\psi^2 \end{bmatrix} \tag{24}$$

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 arbitrary, 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-scan data. By extracting lines from the odd-indexed laser scan data points, corners are identified with the following characteristics: (i) two lines must be nearly perpendicular, (ii) the endpoints of the lines must be within 5cm of each other, and (iii) the line orientations must match the possible wall orientations in the environment.

The relative position measurement is written as a 3-D vector from the sensor to the landmark which is aligned to the sensor frame and projected down to 2-D:

$$\mathbf{z} = \mathbf{\Pi}\mathbf{C}\left(\mathbf{q}\right)\left({}^G\mathbf{p}_{Li} - {}^G\mathbf{p}_S\right) + \mathbf{n}_p, \ \mathbf{\Pi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{25}$$

where ${}^G\mathbf{p}_{Li}$ is the position of the landmark L_i, ${}^G\mathbf{p}_S$ is the position of the sensor (i.e., the estimated state $\mathbf{x}$), and $\mathbf{n}_p$ is the noise in this measurement.

Applying the expectation operator on both sides of (25), we compute the estimated measurement as:

$$\widehat{\mathbf{z}} = \mathbf{\Pi}\mathbf{C}\left(\widehat{\mathbf{q}}\right)\left({}^G\mathbf{p}_{Li} - {}^G\widehat{\mathbf{p}}_S\right) \tag{26}$$

Finally, differentiation of (25), provides the measurement error equation

$$\begin{aligned} \widetilde{\mathbf{z}} \simeq & -\mathbf{\Pi}\mathbf{C}\left(\widehat{\mathbf{q}}\right){}^G\widetilde{\mathbf{p}}_S + \mathbf{\Pi}\lfloor\mathbf{C}\left(\widehat{\mathbf{q}}\right)\left({}^G\mathbf{p}_{L_i} - {}^G\widehat{\mathbf{p}}_S\right)\times\rfloor\delta\boldsymbol{\theta} + \mathbf{n}_p \\ = & \mathbf{H}_{k+1}\tilde{\mathbf{x}} + \mathbf{n} \end{aligned} \tag{27}$$

where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ encompasses both the measurement noise $\mathbf{n}_p$ and the error $\delta\boldsymbol{\theta}$ in the attitude estimate of the cane.

The measurement update requires to compute the measurement residual $\mathbf{r}_{k+1} = \mathbf{z}_{k+1} - \widehat{\mathbf{z}}_{k+1}$ and the Kalman gain

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

The state estimate is updated as

$$\widehat{\mathbf{x}}_{k+1|k+1} = \widehat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\left(\mathbf{z}_{k+1} - \widehat{\mathbf{z}}_{k+1}\right). \tag{28}$$

Lastly, the updated covariance is computed as

$$\begin{aligned} \mathbf{P}_{k+1|k+1} = & \left(\mathbf{I}_{2\times 2} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right)\mathbf{P}_{k+1|k}\left(\mathbf{I}_{2\times 2} - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right)^{\mathrm{T}} \\ & + \mathbf{K}_{k+1}\mathbf{R}\mathbf{K}_{k+1}^{\mathrm{T}} \end{aligned} \tag{29}$$

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 unobtrusive to the user. For this reason two of the three primary sensors are mounted on the white cane (cf. Fig. 3), and the third sensor is foot mounted. These sensors were interfaced to a laptop via USB, RS-232, and bluetooth, respectively. The real-time software components are written in C++, whereas the software for simulation and data plotting is written in Matlab. The sensor bay is mounted beneath a white cane which measures 1.27m when extended and 0.33m when retracted. We have used a light-weight carbon fiber cane 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 5cm by 5cm by 7cm. It has an angular scan range of 240 degrees, with an accuracy of $\pm 1\%$ of the measurement for distances 1m to 4m. Closer than 1m, the measurement accuracy is $\pm$10mm. The laser scanner can measure distances ranging from 0.02m to 4m. The scanner weighs only 160g and consumes 2.5W at 5V. The 3-axis gyroscope is an ISIS Inertial Measurement Unit (IMU), with an angular-rate range of $\pm$90deg/sec. Over an RS-232 connection, the ISIS IMU provides measurements at 100Hz. The weight of the sensor is 363g, and the power consumption is 6.72W at 12V. The IMU measures $5.5\times6.5\times7$cm. The pedometer is $8.5\times3.5\times3.5$cm, and transmits communication packets via bluetooth at a rate of 1Hz.

B. Description of the Experiment

The method described in this paper was tested in an indoor environment on a closed loop of path length 130m.

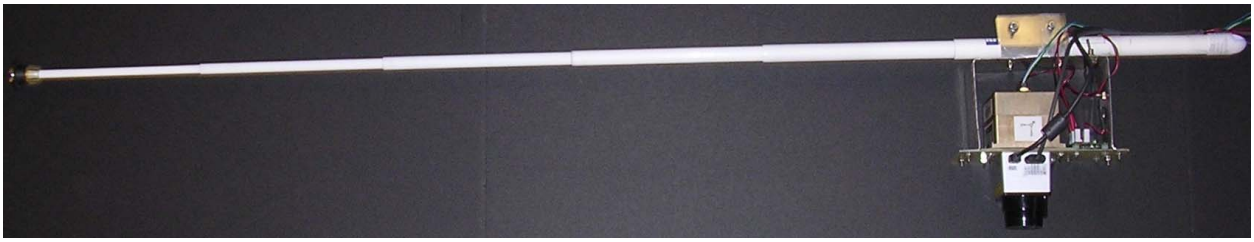

Fig. 3. A view of white cane and hardware. The sensor bay is mounted on the handle portion of the cane. Note that the tip of the cane is unobstructed, and the user maintains the ability to physically sense the world. The weight of the sensor package is approximately 550g.

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-C.2. While walking around, the user testing the cane swung it to-and-fro in a natural manner searching for obstacles which might lie in their path. Fig. 6 shows the estimated trajectory superimposed 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.

Surprisingly, the uncertainty in the position estimate is very low (max $\sigma = 0.16$m), even though the number of position update measurements is small (only 9 of the corners were detected in approximately 110 laser scans). The reason that the position filter is so precise despite the relatively infrequent position update measurements is due to the accuracy of the heading estimates provided by the attitude filter. The attitude estimate of the cane is highly accurate due to over $5,000$ orientation measurements obtained during this experiment (cf. Fig. 1). Based on the analysis of [23] we can infer that when the orientation error is bounded, i.e., $\sigma_\psi \leq \sigma_{\psi_0}$, then the position covariance grows as:

$$\mathbf{P}(t) \simeq 0.5\left(\sigma_v^2 + \sigma_{\psi_0}^2 V^2 \delta t^2\right) \; t = \alpha \; t \qquad (30)$$

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 [22], 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 0.054m^2 which corresponds to approximately 0.16m 1σ error in each direction (cf. Fig. 4). Furthermore, the filter maintains consistency as the x and y components of the measurement residuals fall within the 3σ bounds of the residual covariance (cf. Fig. 5).

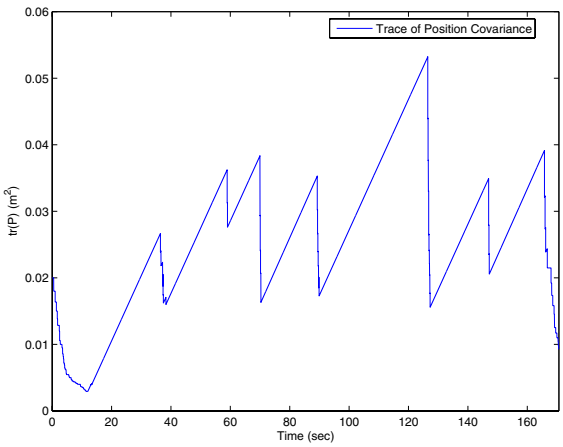

Fig. 4. The trace of the position covariance verifies that the positioning uncertainty remains bounded, but grows linearly with time between consecutive position updates.

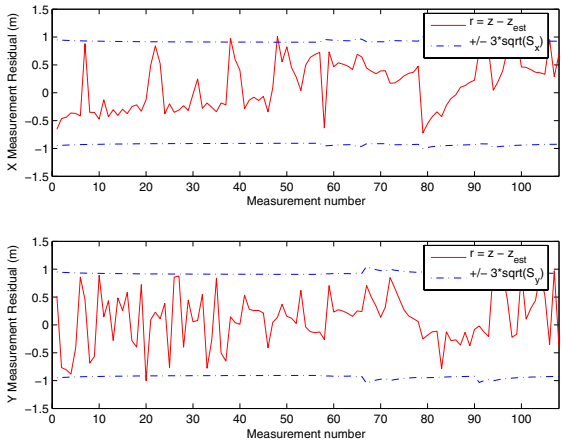

Fig. 5. The x- and y-axis position measurement residuals plotted with their corresponding 3σ bounds computed from the residual covariance.

V. Conclusions and Future Work

This paper has presented a new approach to indoor localization for the visually impaired. Information from a pair of cane-mounted sensors, and a foot-mounted pedometer with unknown and time-varying relative coordinate transformation was fused in a two-stage pose estimator. The first stage utilized inertial measurements from a 3-axis gyroscope and relative orientation measurements from laser scan data to accurately estimate the attitude of the white cane. The second stage estimated the position of the person holding the cane, by incorporating linear velocity measurements from the pedometer, a filtered version of the cane's yaw estimate, and corner features extracted from the laser scan data.

Our estimation algorithms were experimentally validated, and shown to be robust even in a dynamic environment. The accuracy of these results are among the benefits of using a

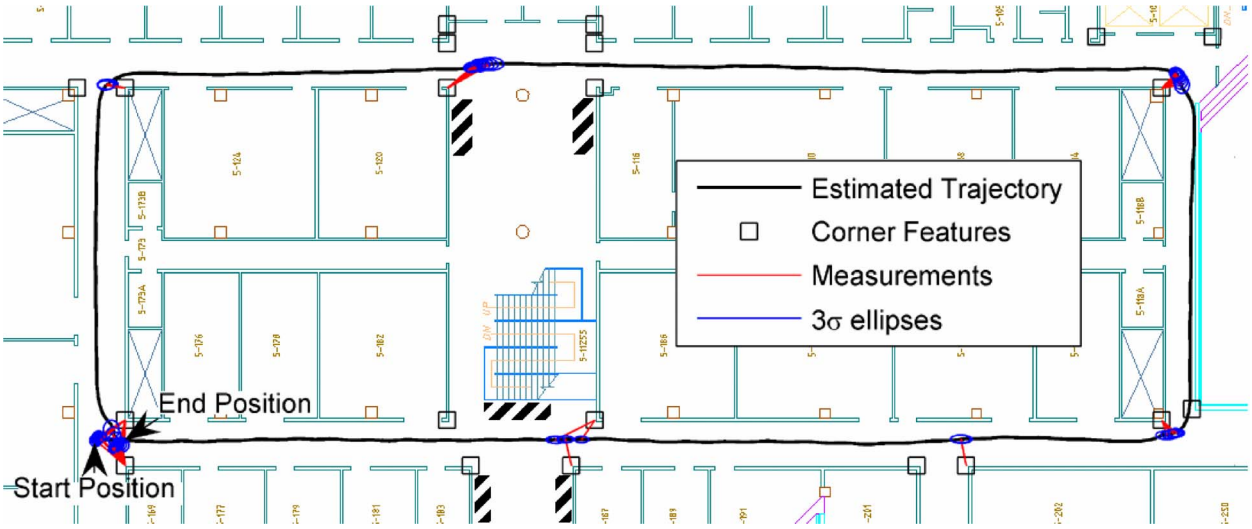

Fig. 6. Experimental results from a 130m test run, starting in the lower left corner, traveling counter-clockwise and ending back at the start location. The striped regions on the figure depict locations of furniture and other objects not represented in the blueprint.

cane-based system for indoor human localization. These also include: (i) a cane provides a lightweight, and unobtrusive platform for sensors, (ii) it is a trusted tool for blind people, and (iii) the user maintains the ability to *physically touch* the environment during operation.

Future work 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 data assistant, or an embedded computer is also within our short term goals. Finally, we intend to expand on this work and also address the case of indoor localization, using the white cane, for cases where a map of the building is not available.

REFERENCES

[1] 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*, vol. 28, pp. 459–467, Aug 1998.

[2] L. Ran, S. Helal, and S. Moore, "Drishti: an integrated indoor/outdoor blind navigation system and service," in *Proc. of the Second IEEE Annual Conference on Pervasive Computing and Communications*, Orlando, FL, 2004, pp. 23 – 30.

[3] 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 Ninth International Conference on Road Transport Information and Control*, London, UK, 1998, pp. 55–58.

[4] D. Yuan and R. Manduchi, "A tool for range sensing and environment discovery for the blind," in *IEEE Workshop on Real-Time 3-D Sensors and Their Use*, 2004.

[5] 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 International Conference on Robotics and Automation*, May 1994, pp. 2012–17.

[6] I. Ulrich and J. Borenstein, "The guidecane - applying mobile robot technologies to assist the visually impaired," *IEEE Transactions on Systems, Man, and Cybernetics, -Part A: Systems and Humans*, vol. 31, no. 2, pp. 131–136, Mar. 2001.

[7] I. Shim and J. Yoon, "A robotic cane based on interactive technology," in *Proc. of the 28th Annual Conf. of the IEEE Industrial Electronics Society (IECON 2002)*, Sevilla, Spain, 2002, pp. 2249–54.

[8] A. Hub, J. Diepstraten, and T. Ertl, "Design and development of an indoor navigation and object identification system for the blind," in *ASSETS'04*, Atlanta, GA, Oct. 18-20 2004.

[9] F. Wong, R. Nagarajan, and S. Yaacob, "Application of stereovision in a navigation aid for blind people," in *Proc. of the 2003 Joint Conf. of the Fourth Int. Conf. on Information, Communications and Signal Processing and Fourth Pacific-Rim Conf. on Multimedia (ICICS-PCM)*, Singapore, 2003, pp. 734–7.

[10] J. J. Leonard and H. F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 376–382, June 1991.

[11] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.

[12] H. Makino, I. Ishii, and M. Nakashizuka, "Development of navigation system for the blind using gps and mobile phone combination," in *Proc. of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Bridging Disciplines for Biomedicine'*, Amsterdam, Netherlands, 1997, pp. 506–7.

[13] W. Balachandran, F. Cecelja, and P. Ptasinski, "A GPS based navigation aid for the blind," in *Proc. of the 17th International Conference on Applied Electromagnetics and Communications (ICECom 2003)*, Dubrovnik, Croatia, 2003, pp. 34–36.

[14] S. Ceranka and M. Niedzwiecki, "Application of particle filtering in navigation system for blind," in *Proc. of the Seventh International Symposium on Signal Processing and its Applications (ISSPA 2003)*, vol. 2, Paris, France, 2003, pp. 495–8.

[15] V. Kulyukin, C. Gharpure, P. Sute, N. D. Graw, and J. Nicholson, "A robotic wayfinding system for the visually impaired," in *Proceedings of the Sixteenth Innovative Applications of Artificial Intelligence Conference (IAAI-04)*, San Jose, CA, 2004.

[16] V. Kulyukin, P. Sute, and N. D. Graw, "Human-robot interaction in a robotic guide for the visually impaired," in *Proc. of the AAAI Spring Symposium on Interaction between Humans and Autonomous Systems over Extended Operation*, Palo Alto, CA, March 2004.

[17] V. Kulyukin, C. Gharpure, J. Nicholson, and S. Pavithran, "Rfid in robot-assisted indoor navigation for the visually impaired," in *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, Sep. 28-Oct. 2 2004.

[18] S. Ertan, C. Lee, A. Willets, H. Tan, and A. Pentland, "A wearable haptic navigation guidance system," in *Proc. Second Int. Symposium on Wearable Computers*, Pittsburgh, PA, 1998, pp. 164–165.

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

[20] N. Trawny and S. I. Roumeliotis, "Indirect Kalman filter for 6D pose estimation," University of Minnesota, Dept. of Comp. Sci. & Eng., Tech. Rep. 2005-002, Jan. 2005.

[21] P. S. Maybeck, *Stochastic models, estimation, and control*, R. Bellman, Ed. Academic Press, 1979, vol. 1.

[22] A. Kelly, "Linearized error propagation in odometry," *International Journal of Robotics Research*, vol. 23, no. 2, pp. 179–218, Feb. 2004.

[23] A. I. Mourikis and S. I. Roumeliotis, "Performance bounds for cooperative simultaneous localization and mapping (c-slam)," in *Robotics: Science and Systems*, Boston, MA, 2005.