FootSLAM: Pedestrian Simultaneous Localization and Mapping Without Exteroceptive Sensors—Hitchhiking on Human Perception and Cognition

The authors describe a Bayesian estimation approach that achieves simultaneous localization and mapping for pedestrians using odometry obtained with foot-mounted inertial sensors.

By Michael Angermann, *Member IEEE*, and Patrick Robertson, *Member IEEE*

ABSTRACT | In this paper, we describe FootSLAM, a Bayesian estimation approach that achieves simultaneous localization and mapping for pedestrians. FootSLAM uses odometry obtained with foot-mounted inertial sensors. Whereas existing approaches to infrastructure-less pedestrian position determination are either subject to unbounded growth of positioning error, or require either *a priori* map information, or exteroceptive sensors, such as cameras or light detection and ranging (LIDARs), FootSLAM achieves long-term error stability solely based on inertial sensor measurements. An analysis of the problem based on a dynamic Bayesian network (DBN) model reveals that this surprising result becomes possible by effectively hitchhiking on human perception and cognition. Two extensions to FootSLAM, namely, PlaceSLAM, for incorporating additional measurements or user provided hints, and FeetSLAM, for automated collaborative mapping, are discussed. Experimental data that validate FootSLAM and its extensions are presented. It is foreseeable that the sensors and processing power of future devices such as smartphones are likely to suffice to position the bearer with the same accuracy that FootSLAM achieves with foot-mounted sensors already today.

KEYWORDS | FeetSLAM; FootSLAM; odometry; pedestrian navigation; simultaneous localization and mapping (SLAM)

I. INTRODUCTION

FootSLAM is an algorithm that apparently challenges a well-established conjecture in navigation: "Navigation based on inertial sensors as sole means is subject to unbounded growth of position error over time." This unbounded growth of position error has very practical implications, especially in mass market pedestrian navigation, where straightforward dead reckoning with low-cost inertial sensors results in hundreds of meters of error after a few minutes. If the navigation makes explicit use of the

human legged locomotion by detecting steps, inertial sensors can be recalibrated during the rest phases of the foot, which results in significant reduction of the error growth to say a few meters after minutes. Since the error is still a random walk, even if this technique is applied, it typically results in hundreds of meters of error after few hours. In contrast, we have shown in experiments that FootSLAM is able to maintain an error in the range of one meter in typical indoor environments over extended periods of time without noticeable growth of error. As we will explain in the course of this paper, FootSLAM actually does not violate the conjecture stated above. While the inertial sensors are the "sole (technical) means," FootSLAM effectively augments the inertial sensors with other sensors, namely the human's perception. In dead reckoning, a current position estimate is computed from the previous estimate and the estimated relative movement (odometry). Given an ideal inertial sensor, with which the relative movement can be measured without error, the position can be determined for all times with the accuracy of the initial position. For nonideal sensors the errors are cumulative and the growth in uncertainty is unbounded over time. FootSLAM builds on prior work on pedestrian positioning using foot-mounted inertial measurement units (IMUs) and the simultaneous localization and mapping (SLAM) approach pioneered in robotics [1], [2]. The novelty is that FootSLAM uses no visual or other exteroceptive sensors of any kind (the only sensors are the accelerometers and gyroscopes of a foot-mounted IMU) and still is able to prevent unbounded error growth. While this is still an unproven conjecture from a theoretical standpoint, we were able to show in our experiments that a pedestrian's location and the building layout can be jointly estimated by using the pedestrian's IMU-based odometry alone.

A. Motivation

Moving as a pedestrian constitutes the most ubiquitous element of transportation in human society. The efficient routing of pedestrians in public transportation hubs is a prerequisite for tightly coupled and efficient intermodal transportation. Personal navigation aids require location and map accuracy in the order of the physical extension of relevant structures (e.g., doors, stairs, aisles) in the environment which is typically around one meter for indoor environments. Hence, the motivation for our work is to enable location determination and mapping for pedestrians in these environments where other means of location determination, such as satellite navigation or other radio-based localization techniques or mapping information are often not available or too inaccurate. While this has been achieved for robotic platforms in the past, it has been an unsolved problem for pedestrians primarily for one reason: exteroceptive sensors, such as cameras or light detection and ranging (LIDARs) are standard for robots, but problematic for pedestrian applications due to their

mounting requirements, cost, and privacy concerns. Simultaneous localization and mapping for pedestrians, based solely on IMU data, would remove significant hurdles for mass market and professional applications, by eliminating the need for exteroceptive sensors and reducing privacy issues.

Projecting the improvements in microelectromechanical system (MEMS)-based inertial sensors to the next ten or even 20 years, it is foreseeable that FootSLAM will no longer require the inertial sensors to be mounted on the foot. Instead, the sensors and processing power in a pocket-carried smartphone are likely to suffice to position the bearer with the same accuracy that FootSLAM achieves with foot-mounted sensors already today. In essence, users of ordinary smartphones would always know their position and simultaneously create a map of their environment, merely by walking around in it.

B. Related Work

Automated dead reckoning has a long tradition in naval, aviation, and automotive position determination. Its application to automated determination of a pedestrian's position has been facilitated only recently by the advent of highly integrated MEMS-based IMUs. Specifically, its application in mass market applications has become commercially prospective by the development of low-cost silicon turn rate sensors [3]. In order to cope with these sensors' errors in terms of time-varying biases and scale factors, Dissanayake et al. employed movement constraints based on a vehicle model within a Bayesian filter [4]. While their application domain was land vehicles, they effectively laid the foundation for online estimation of IMU sensor errors by using movement constraints which would later become the central pillar of inertial pedestrian position determination. Early work on dead reckoning for pedestrians relied on simple step detection and assumed either a fixed step length or a correlation of step frequency and step length. While these methods do not employ a strapdown algorithm to solve for a full 3-D displacement vector of the foot, they are fairly robust and, typically in conjunction with an electronic compass, work also for sensors mounted on the hip or upper torso. Randell et al. improved the estimation of the step length by measuring the peak acceleration of the foot in each step and using it as additional input [5]. The shortcoming of these methods is their inability to correctly determine nontrivial steps, such as sidestepping. Foxlin recognized the possibility to detect and use the duration of the foot's rest phase for injecting zero-velocity updates (ZUPTs) as pseudomeasurements into a filter that models the IMU's errors [6] (Fig. 1). Closely related to utilizing a platform's dynamic constraints, which may arise from its inertia or limited power and are typically constant over time and location, is the idea to exploit constraints of movement imposed by the environment. Such constraints may be caused by walls or other obstacles and are typically a function of the
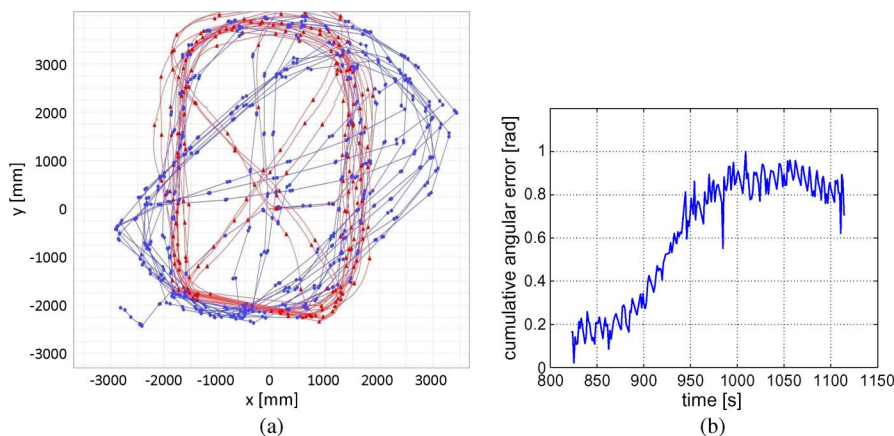
Fig. 1. *Track of IMU-based odometry (blue) after applying ZUPT pseudomeasurements compared to ground truth (red) and temporal evolution of the cumulative angular error of the odometry. At $t \approx 870$ s the angular error starts to grow with approximately constant rate until $t \approx 1000$ s. Data from data set 16 of [12], with AM1T3ND ZUPT. (a) Track. (b) Cumulative angular error.*

location. By using known building layouts, several groups achieved long-term stability in position determination of a pedestrian in 2-D [7], [8] and 2.5-dimensional [9] environments. Common to their approaches is the use of nonparametric sequential Monte Carlo filters ("particle filters"). Given the prerequisite of a sufficiently constraining environment and an accurate map, these methods achieve long-term error stability. A huge body of work exists in the robotics literature on the use of a wide range of sensors, such as sonar, laser ranging, and cameras to perform positioning of robots. The simultaneous localization and mapping (SLAM) problem was formulated to allow robots to navigate in *a priori* unknown environments [1]: a moving robot explores its environment and uses its sensor information to build a "map" of landmarks. Our work is closely related to the Rao-Blackwellized particle filtering approach employed in the FastSLAM algorithm [10]. Furthermore, we employ a probabilistic map that represents human motion in a 2-D hexagonal grid that is similar to an occupancy grid [11] but with a different purpose.

C. Problem Statement

Existing approaches to infrastructure-less pedestrian position determination are either subject to unbounded growth of positioning error, or require one of the following two: 1) *a priori* map information, or 2) exteroceptive sensors, such as cameras or LIDARs with which traditional SLAM is being performed. In contrast, we wish to achieve long-term stability of pedestrian positioning, i.e., bounding the positioning error based solely on nonperfect odometry that exhibits both angular and distance errors.

D. Structure of the Paper

After briefly discussing the motivation, related work, and our problem statement, we first present the Bayesian formulation of FootSLAM in Section II. The original FootSLAM algorithm and its extensions PlaceSLAM and FeetSLAM are outlined in Section III. Section IV describes experiments performed to validate the approach, and presents and discusses results from these experiments. The paper closes with conclusions and an outlook in Section V.

II. BAYESIAN DERIVATION

A. Dynamic Bayesian Network Representation

We will formulate the problem as a dynamic Bayesian network (DBN). The key is to suitably represent the actor, i.e., the pedestrian in the system. When a pedestrian walks in a constrained environment, he relies mainly on visual cues in order to avoid walls and other obstacles. The pedestrian might be walking toward a particular destination such as an office, or might just be walking randomly in the accessible space in an office during a conversation. In robotic SLAM, the robot's movement is controlled by a series of inputs $u(t)$. These inputs are then used in the SLAM estimation as inputs to a probabilistic motion model. For FootSLAM, we assume that the human visual and cognitive systems interpret the environment and use it to guide motion: observed physical constraints such as walls influence intentions which result in a person deciding which steps to take. Fig. 2(b) shows a DBN that models relevant aspects of the system. All random variables are denoted in bold face. The step transition vector $\mathbf{U}_k$ has a special property: given the old and new poses $\mathbf{P}_{k-1}$ and $\mathbf{P}_k$, the step transition $\mathbf{U}_k$ is determined entirely, since knowledge of any two of the state variables $\mathbf{P}_{k-1}$, $\mathbf{P}_k$, and $\mathbf{U}_k$ determines the third. Inspecting the DBN, we can now ask which random variables might be measurable or indirectly observable. It has been shown
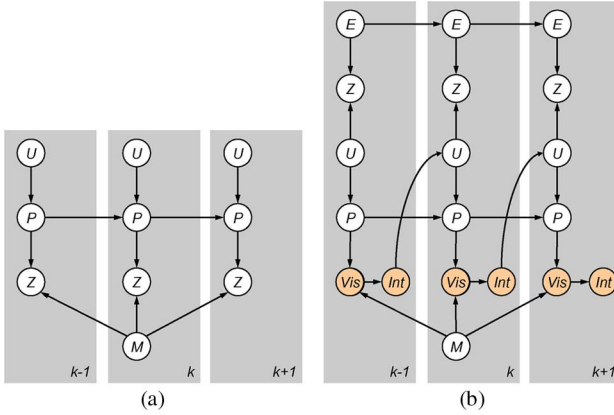
Fig. 2. *DBNs for classic (robotic) SLAM and FootSLAM showing three time slices, with random variables M ("map"), Z ("measurement"), P ("pose"), and U ("odometry"). The FootSLAM map implicitly encodes the environmental features that influence the pedestrian's visual impression (Vis) and intention (Int). E models the correlated errors of the step estimator. (a) Classic SLAM. (b) FootSLAM.*

that observing human visual sensory input is possible, as reported in [13] and more recently for dynamic visual input in [14].[1] However, we assume to have no means of directly observing the human visual system **Vis**, nor can we directly measure where the person might actually want to go next (**Int**), even though step estimation by measuring electrical activity produced in skeletal muscles by electromyography (EMG) has been reported in [15].

In FootSLAM, we make no assumptions as to how steps are measured, as long as the error processes can be modeled sufficiently. So far we have used inertial sensors, in particular, a foot-mounted IMU, that (differentially) measure the steps a person takes. For this case, step measurements $\mathbf{Z}_k^U$ are obtained in a manner as described in [8] and we assume that a suitable strapdown inertial navigation algorithm with a Kalman filter or similar algorithm is used. From the viewpoint of our DBN this estimate is a noisy measurement of the true step vector $\mathbf{U}_k$. The only other influence on the measurement $\mathbf{Z}_k^U$ is a state variable $\mathbf{E}_k$ that encodes the correlated errors of the step estimator.

B. Sketch of Derivation

Our goal is to estimate the states and state histories of the DBN given the series of all observations $\mathbf{Z}_{1:k}^U$ from the foot-mounted IMU. More formally, we wish to compute

the joint posterior $p(\mathbf{P}_{0:k}\mathbf{U}_{0:k}\mathbf{E}_{0:k}, \mathbf{M}|\mathbf{Z}_{1:k}^U)$, which is factorizable

$$
\begin{aligned}
p\big(\mathbf{P}_{0:k}\mathbf{U}_{0:k}\mathbf{E}_{0:k}, \mathbf{M}|\mathbf{Z}_{1:k}^U\big) \\
= p(\mathbf{M}|\mathbf{P}_{0:k}) \cdot p\big(\{\mathbf{P}\ \mathbf{U}\ \mathbf{E}\}_{0:k}|\mathbf{Z}_{1:k}^U\big). \quad (1)
\end{aligned}
$$

The expression for the map probability in (1) simplifies because the assumed knowledge of $\mathbf{P}_{0:k}$ makes the map $\mathbf{M}$ conditionally independent of $\mathbf{U}_{0:k}$, $\mathbf{E}_{0:k}$, and the measurements $\mathbf{Z}_{1:k}^U$; as follows from the DBN in Fig. 2(b) and the deterministic relationship linking $\mathbf{P}_{k-1}$, $\mathbf{P}_k$, and $\mathbf{U}_k$. We will express the second factor in (1) recursively in the sense of a Bayesian filter. It can be easily shown that the recursive formulation is $p(\{\mathbf{P}\ \mathbf{U}\ \mathbf{E}\}_{0:k}|\mathbf{Z}_{1:k}^U)$

$$
\begin{aligned}
\propto \quad & p\big(\mathbf{Z}_k^U|\mathbf{U}_k\mathbf{E}_k\big) \cdot p\big(\{\mathbf{PU}\}_k|\{\mathbf{P}\ \mathbf{U}\}_{0:k-1}\big) \\
& \cdot p\big(\mathbf{E}_k|\mathbf{E}_{k-1}\big) \cdot p\big(\{\mathbf{P}\ \mathbf{U}\ \mathbf{E}\}_{0:k-1}|\mathbf{Z}_{1:k-1}^U\big). \quad (2)
\end{aligned}
$$

The recursion is usually begun with the pose $\mathbf{P}_0$ set to an arbitrary position and heading, since performing SLAM without any absolute heading or location information is invariant over rotation and translation. We assume pertinent postprocessing to resolve rotation, translation (and scale) transformations.

It is clear from the DBN that the map must play a role in determining the second factor of (2), the pose and step transition probability. Marginalizing over $\mathbf{M}$, we write this factor as

$$
I \triangleq \int_{\mathbf{M}} p\big(\{\mathbf{PU}\}_k|\mathbf{P}_{k-1}, \mathbf{M}\big) \cdot p(\mathbf{M}|\mathbf{P}_{0:k-1})\, d\mathbf{M}. \quad (3)
$$

If we are able to compute $p(\mathbf{M}|\mathbf{P}_{0:k-1})$ and the influence of a map on the transition from $\mathbf{P}_{k-1}$ to $\mathbf{P}_k$, we can perform sequential Bayesian estimation of the map and pose.

III. ORIGINAL ALGORITHM AND EXTENSIONS

A. FootSLAM

In order to obtain an online particle filter, we sample from a proposal density, such as the one described in [16], and correct the weights using importance sampling. Following [17] and [18], it can be shown that the weight of the ith particle $w_k^i \propto w_{k-1}^i \cdot I^i$ to a very good approximation. We therefore apply a Rao-Blackwellized particle filter [10] based on (1): each particle i represents $\{\{\mathbf{PUE}\}_k^i, p(\mathbf{M}|\mathbf{P}_{0:k}^{\ i})\}$.

---

[1]A tempting hypothetical experiment would be to record blood oxygen level-dependent (BOLD) signals of the occipitotemporal visual cortex, and then to investigate if Bayesian estimation of the location of the person or the map of the environment is possible. To validate this in absence of portable functional magnetic resonance imaging (fMRI) one could draw on video playback to a subject in a stationary fMRI device.

1) Probabilistic Transition Map: We now introduce a probabilistic map, based on the probability of the pedestrian crossing transitions in a regular 2-D grid of adjacent hexagons of radius r. We choose hexagons because they are the polygons with the greatest number of edges that can be arranged to cover a 2-D area completely without overlap. Furthermore, six angular transitions appear an appropriate number of choices for which human motion would be reasonably independent between angles.

We restrict this space to the region visited by any particle and define H_h as one of N_H hexagons, where h uniquely references a hexagon's position. We define the map $\mathbf{M}$ to be the set comprising all $\mathbf{M}_h$, where $\mathbf{M}_h$ is a vector of length 6 with each component denoting the transition probability

$$\mathbf{M}_{h(\mathbf{P}_{k-1})}^{e(\mathbf{U}_k)} \stackrel{\triangle}{=} P(\mathbf{P}_k \in H_j | \mathbf{P}_{k-1} \in H_h), \quad \text{where } j \neq h \quad (4)$$

for leaving the hth hexagon over the edge e via $\mathbf{U}_k$ to an adjacent hexagon H_j. We assume that the map factors into local, conditionally independent components, and step $\mathbf{U}_k$ is only dependent on $\mathbf{P}_{k-1}$ and the local map $\mathbf{M}_{h(\mathbf{P}_{k-1})}$. Writing $\tilde{h}$ for outgoing hexagon $h(\mathbf{P}_{k-1})$, and $\tilde{e}$ for $e(\mathbf{U}_k)$, we compute the integral I by integrating over $\mathbf{M}_{\tilde{h}}^{\tilde{e}}$ for the respective edge.

2) Learning the Transition Map: Learning is straightforwardly based on Bayesian learning of probabilities of discrete random variables. Each time a particle makes a transition $\mathbf{P}_{k-1}^i \rightarrow \mathbf{P}_k^i$ across edge $\tilde{e}$ we count this transition in its local map of hexagon $H_{\tilde{h}}$. When computing the counts, we assume that observing a certain transition from an outgoing hexagon to an incoming one allows us to increment the counts for both the outgoing as well as the incoming one. This improves convergence and is the same as assuming that a person is likely to walk in either of the two directions. In order to incorporate prior information, we assume that $p(\mathbf{M}_{\tilde{h}}^{\tilde{e}} | \mathbf{P}_{0:k})$ follows a beta distribution and integrating I^i yields

$$I^i \propto \frac{N_{\tilde{h}}^{\tilde{e}} + \alpha_{\tilde{h}}^{\tilde{e}}}{N_{\tilde{h}} + \alpha_{\tilde{h}}} \quad (5)$$

where $N_{\tilde{h}}^{\tilde{e}}$ is the number of times the ith particle crossed the transition, $N_{\tilde{h}}$ is the sum of the counts over all edges of the hexagon in this particle's map counters, and $\alpha_{\tilde{h}}^{\tilde{e}}$ and $\alpha_{\tilde{h}} = \sum_{e=0}^{5} \alpha_{\tilde{h}}^e$ are the prior counts. So far an increment of the time index k is associated with a step that leads from one hexagon to an adjacent one. In reality, a step might keep the hypothesized pose in the hexagon or it might lead it over several hexagons. To address this we simply perform a weight update only when hypothesized pose has stepped out of the

last hexagon and apply multiple products of (5) for all edges crossed. Similarly, we update the counts of all edges crossed.

Algorithm 1: FootSLAM

1: **for** $i = 1 \rightarrow N_p$ **do**
2:    $\mathbf{P}_0^i \leftarrow (x, y, h = 0)$, where x, y, h denote the pose location and heading in 2-D
3:    $\mathbf{E}_0^i \leftarrow$ draw from initial distr. of odometry error states
4: **end for**
5: **for all** time steps **do**
6:    $k \leftarrow k + 1$
7:    **for** $i = 1 \rightarrow N_p$ **do**
8:       draw the ith particle from the proposal density $p(\mathbf{E}_k | \mathbf{E}_{k-1}^i) \cdot p(\mathbf{U}_k | \mathbf{Z}_k^U, \mathbf{E}_k^i)$ from left to right.
9:       $\mathbf{P}_k^i \leftarrow \mathbf{P}_{k-1}^i + \mathbf{U}_k^i$
10:      $w_k^i \leftarrow w_{k-1}^i \cdot I_k^i$
11:      where $I_k^i \propto \prod_{\forall \text{ edges}} (N_{\tilde{h}}^{\tilde{e}i} + \alpha_{\tilde{h}}^{\tilde{e}} / N_{\tilde{h}}^i + \alpha_{\tilde{h}})$
12:      **for all** edges crossed **do**
13:        **for** both hexagons joined by the edge $\tilde{e}^i$ **do**
14:          $N_{\tilde{h}}^{\tilde{e}i} \leftarrow N_{\tilde{h}}^{\tilde{e}i} + 1$ w.r.t. the outgoing hexagon $\tilde{h}^i$
15:        **end for**
16:      **end for**
17:    **end for**
18:    normalize so that $\sum_{i=1}^{N_p} w^i = 1$
19:    **if** resampling is required **then**
20:      perform resampling
21:    **end if**
22: **end for**

B. PlaceSLAM

A straightforward extension of FootSLAM is the detection of physical proximity to reliably detectable "places" [19]. This detection hints may either stem from an assisting user or additional sensors, such as a radio-frequency identification (RFID)-tag reader or a camera that recognizes salient visual features. It is important to recognize that PlaceSLAM does not require continuous measurements, but can make use of sporadically incoming place detection incidents. A user may choose suitable places and hint their proximity when passing them. In subsequent walks, the person may signal when revisiting these places. Cues can be prominent items such as fire extinguishers or anything that would allow reliable and repeatable recognition of location.

Depending on the distinguishability of the hints or features, different variants of PlaceSLAM exist. In a trivial case, the true locations of the places are known to the system. In less trivial cases, the locations of placemarks are not known and the association "quality" of place identifiers may range from perfect *a priori* association to no *a priori* association at all. Fig. 3 shows a stylized trajectory of a pedestrian. Circles represent places. Letters and colors
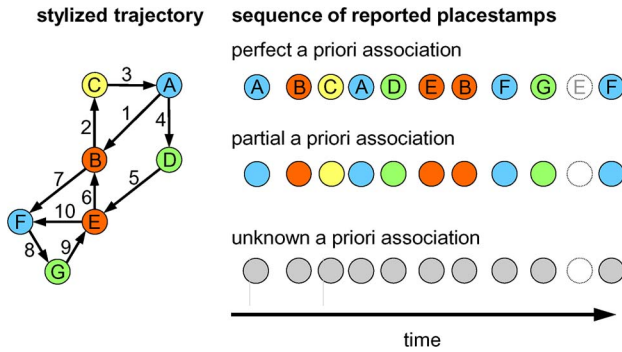
Fig. 3. *Stylized trajectory of a pedestrian. Reporting of placestamps may be incomplete (e.g., "E" is missing after leg 9).*

are identifiers. On the right side of the figure we see three possible placestamp sequences as input to the estimator: one for which the placestamps carry the unique letters, i.e., perfect association; one with partial association (colors); and finally, one in which only the fact that some place is seen is reported, i.e., unknown association. Note that we do not require uniqueness of the identifier. One can imagine a situation where a pedestrian signals every time he walks through any door. The requirement in this case is that such places are sufficiently separated in space. This is the most challenging and general case for PlaceSLAM.

For each time step, the standard FootSLAM proposal step and weight update is performed for each particle. If no placestamp has been reported, we continue with the normal FootSLAM algorithm. If a placestamp was reported, we distinguish two cases. 1) If the particle position is separated wider than a predefined threshold distance $d_{\min}$ from all previously logged places in the particle's own place map, we assign a new unique identifier to this place. 2) Otherwise, we select the identifier of the place in the particle's place map closest to the particle's current position. In both cases, we then weight the particle with the product of the FootSLAM weight and the PlaceSLAM weight. In both cases, we subsequently update the location of the place according to its previous location distribution and to the location of the particle [19].

C. FeetSLAM

If multiple data sets, possibly from multiple pedestrians, are combined, both coverage and accuracy can be increased. FeetSLAM, an extension to FootSLAM, combines multiple odometry data sets in an iterative process [20]. The result of each iteration enters the next iteration through (5) in the form of prior information. This approach is similar and in fact inspired by the iterative "Turbo" decoding principle applied in communication systems [21], which in turn is an instance of loopy belief propagation [22], [23]. Since the iterative process allows to combine the resulting data sets from walks of many dif-

ferent pedestrians with tractable complexity, FeetSLAM enables the computation of community-generated maps. In principle, these community-generated maps could form a map database with global coverage.

IV. EXPERIMENTAL VALIDATION

In order to investigate the conjecture that FootSLAM is able to bound the growth of positioning error, we carried out initial experiments in an office building environment. A pedestrian was instrumented with a foot-mounted IMU and performed three walks, each of roughly 10-min duration, within this environment. We recorded ground truth for two positions at opposite corners of the main corridor by timestamping the event everytime the pedestrian passed these positions. All data sets were recorded and processed offline. A similar data set with additional place-marks has been recorded in the same environment. A detailed analysis of the results can be found in [19]. FeetSLAM has been tried in larger and more complex environments. Fig. 4 shows a map of an environment that included straight and curved corridors, nonrectangular crossings, and several loops [20].
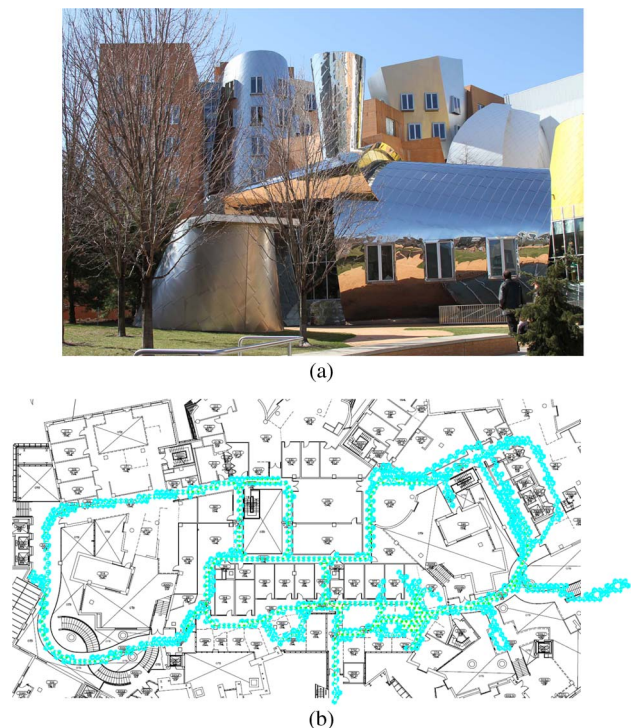

(a)

(b)

Fig. 4. *Inertial measurements of several partially overlapping walks are automatically combined by FeetSLAM, a collaborative form of FootSLAM, to estimate the map shown above. (a) MIT's Stata Center, a large academic building with many nonrectangular structures. (b) Maximum a posteriori (MAP) FeetSLAM map (cyan) resulting from several walks on one floor of Stata Center.*

A. Discussion of Results

1) Performance: A learned map (manually translated, rotated, and scaled) for hexagon radius $r = 0.5$ m is shown in Fig. 5. In the visited areas, this map reflects the real path and is accurate to about 1–3 m, with better accuracy in the corridors that were frequented more often. The walk intentionally remained out of "loop closure" in the corridor for some time. The particles start to converge once the user backtracks or revisits a region for about 10 m. A sufficient number of hexagons have to be revisited once or twice for a usable map to emerge. This governs the required duration of a walk and fits with what a person typically covers in an office day. Accuracy in any case is related to the physical structure dimension, such as corridors and doors, which is about 1–2 m. The error evolution for FootSLAM is shown in Fig. 6. Our coordinate system origin was both the starting point and one of the reference points and we manually corrected for rotation ambiguity. With a sufficient number of particles we achieve an accuracy of approximately 2 m at the two reference points. Without FootSLAM, we see unbounded error growth after some time—our system coasted without too much error for about 300 s. Durations from 30 to 300 s are typical and suggest that without maps the particle filter can bridge areas like large halls where there are no features for FootSLAM to map. We expect FootSLAM to require a certain minimum average restrictedness of motion. Nevertheless, it can coast over some open areas given enough particles. To achieve accurate mapping, a relatively large number of particles ($> 10\,000$) is necessary. Our current implementation runs roughly in real time for about 30 000 particles on a standard personal
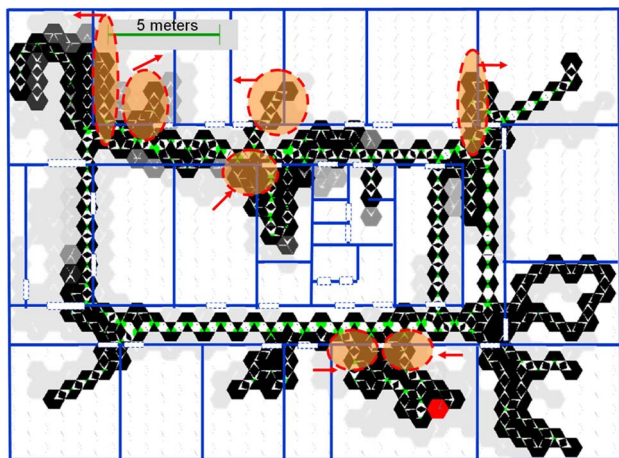

Fig. 6. *Relative position accuracy evaluation for FootSLAM—each curve is a single run on one data set.*

Fig. 5. *A map resulting from FootSLAM based on IMU data alone. Shown is an overlap of the posterior (i.e., weighed average) map in shades of gray and the maximum a posteriori (MAP) map (i.e., the map of the "best" particle) in black. Ovals highlight errors with arrows indicating roughly the error vector. The reference building layout is plotted for comparison.*
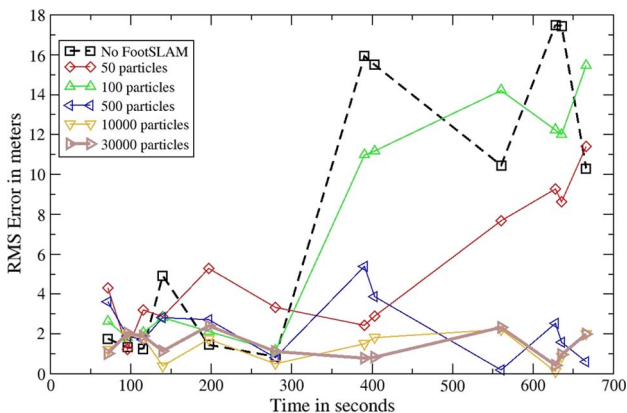
computer (PC). In our proposal function, we drew $\mathbf{E}_k^i$ from two independent random walk processes to model heading bias and heading rate bias error states [16]. The additive error in 2-D space between $\mathbf{Z}_k^U$ rotated according to $\mathbf{E}_k$, and $\mathbf{U}_k$, was white and Gaussian. In addition to SLAM-inherent rotation and translation invariance, FootSLAM is subject to a map scaling error. In FootSLAM, this error is due to biases in the IMU sensors, occasional erroneous ZUPTs, and subsampling/clipping of the IMU signals that affect the step length estimation, and also a result of particles exploring hypotheses of different lengths. In our quantitative evaluation and videos [24], we have *not* used an individual length correction factor. This was only done using a constant factor of 1.15, which we have established quite reliably for that sensor setup. For the illustrations in the known floor plans, we adjusted the scaling so that the resulting map fits the known floor plan. This additional scaling adjustment was less than 10%. Scaling can be automated when tracks are anchored to outdoors Global Positioning System (GPS) measurements or other absolute positioning systems.

2) Privacy Implications: Given the current performance of FootSLAM and projecting the improvements in MEMS-based inertial sensors to the next ten, or even 20 years, a range of significant privacy implications for practically everybody arises, the solution of which is beyond the scope of this paper. We mention them for completeness and believe that they may require some form of regulation or even legislature in the coming years. Whoever has access to the stream of IMU data can determine the pedestrian's current position and the historic trajectory, as well as the layout of the environment. Since FootSLAM can position the user without any external signal or interaction with infrastructure, all sensing, computation, and user interaction can be carried out on the personal device, theoretically with full control over all data by the user. For this

reason, we suggest that data from inertial sensors should be subject to the same access constraints as other location sensors such as GPS.

Perhaps the most unclear privacy or security threat that is caused by the existence of FootSLAM is virtually impossible to prevent or even detect. Covert mapping of environments such as private, industrial, or governmental premises is hard to counter or detect, since FootSLAMs rids the need for any technical exteroceptive sensor that could be found in a nondestructive inspection. On the other hand, FootSLAM may—in terms of privacy—be the acceptable compromise for indoor navigation on such premises, since the IMU data are most likely to be considered less sensitive than the images taken by a pedestrian-worn camera in visual SLAM approaches for pedestrian navigation.

V. CONCLUSION AND OUTLOOK

As of today, FootSLAM is capable of achieving stable position determination and accurate map generation solely based on an IMU in environments without any *a priori* knowledge. Extensions such as PlaceSLAM and FeetSLAM have been introduced and have been shown to work in a wide range of realistic experiments.

We foresee several interesting avenues for further research that will extend this work. So far we have concentrated on feasibility without optimizing for computational complexity. We see the main potential of FootSLAM in its collaborative form (FeetSLAM). In order to operate FeetSLAM in an economically viable way on a global scale, and in a 3-D environment, reducing required memory and computational effort is desirable. This holds both for a centralized approach in which most of the computation is performed on large-scale server farms as well as in a decentralized approach in which data and computation

are shared among mobile devices. Any improvement in the underlying odometry not only improves the accuracy but also reduces computational effort in FootSLAM. Since significant improvements with respect to the accuracy of MEMS-based inertial sensors are to be expected in the next ten years, the future of inertial-based pedestrian navigation holds promise. In the not too distant future, it should be possible to reduce the dependency on "naïve" ZUPTs, which require a foot-mounted IMU. This will allow for less restrictive mounting on the hip or upper body. Ideally, the user will be freed from any constraints on where to put the device, and future algorithms will be able to robustly estimate the sensor's trajectory, independently of whether it is attached to the belt or just loosely put into a pocket as part of an arbitrary mobile device. Should cold-atom interferometry [25], which has already been applied in proof of concepts for inertial sensors with up to six axes [26], one day lead to practical devices, localization will require very little additional information for estimating and correcting the remaining errors. A FootSLAM-like approach would likely suffice to provide this additional information and do so with very little computational effort. Apart from these technical issues, interesting questions arise concerning ownership of data: if Alice walks through Carol's home or company, is she entitled to use FootSLAM to generate a map, use it herself, share it with Bob, or even publish it? ■

REFERENCES

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

[2] J. Leonard and H. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," in *Proc. IEEE Int. Workshop Intell. Robots Syst.*, Osaka, Japan, 1991, Nov. 1991, vol. 3, pp. 1442–1447.

[3] R. Hulsing, "MEMS inertial rate and acceleration sensor," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 13, no. 11, pp. 17–23, Nov. 1998.

[4] G. Dissanayake, S. Sukkarieh, E. Nebot, and H. Durrant-Whyte, "The aiding of a low-cost strapdown inertial measurement unit using vehicle model constraints for land vehicle applications," *IEEE Trans. Robot. Autom.*, vol. 17, no. 5, pp. 731–747, Oct. 2001.

[5] C. Randell, C. Djiallis, and H. Muller, "Personal position measurement using dead reckoning," in *Proc. 7th IEEE*

Int. Symp. Wearable Comput., Los Alamitos, CA, Oct. 2003, pp. 166–173.

[6] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," *IEEE Comput. Graph. Appl.*, vol. 25, no. 6, pp. 38–46, Nov. 2005.

[7] S. Beauregard, Widyawan, and M. Klepal, "Indoor PDR performance enhancement using minimal map information and particle filters," in *Proc. IEEE/ION Position Location Navig. Symp.*, Monterey, CA, May 2008, pp. 141–147.

[8] B. Krach and P. Robertson, "Cascaded estimation architecture for integration of foot-mounted inertial sensors," in *Proc. IEEE/ION Position Location Navig. Symp.*, Monterey, CA, May 2008, pp. 112–119.

[9] O. Woodman and R. Harle, "Pedestrian localisation for indoor environments," in *Proc. 11th ACM Int. Conf. Ubiquitous Comput.*, Seoul, Korea, 2008, pp. 114–123.

[10] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: A factored solution to the simultaneous localization and mapping

problem," in *Proc. AAAI Nat. Conf. Artif. Intell.*, Edmonton, ON, Canada, 2002, pp. 593–598.

[11] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *IEEE Comput. Mag.*, vol. 22, no. 6, pp. 46–57, Jun. 1989.

[12] M. Angermann, P. Robertson, T. Kemptner, and M. Khider, "A high precision reference data set for pedestrian navigation using foot-mounted inertial sensors," in *Proc. IEEE Int. Conf. Indoor Positioning Indoor Navig.*, Sep. 2010, DOI: 10.1109/IPIN.2010. 5646839.

[13] K. Kay, T. Naselaris, R. Prenger, and J. Gallant, "Identifying natural images from human brain activity," *Nature*, vol. 452, pp. 352–355, Mar. 2008.

[14] S. Nishimoto, A. T. Vu, T. Naselaris, Y. Benjamini, B. Yu, and J. L. Gallant, "Reconstructing visual experiences from brain activity evoked by natural movies," *Current Biol.*, vol. 21, no. 19, pp. 1641–1646, Oct. 2011.

[15] R. Chen, W. Chen, X. Chen, X. Zhang, and Y. Chen, "Sensing strides using EMG signal for pedestrian navigation," *GPS Solutions*, vol. 15, pp. 161–170, 2011.

[16] P. Robertson, M. Angermann, B. Krach, and M. Khider, "SLAM dance: Inertial-based joint mapping and positioning for pedestrian navigation," *InsideGNSS Mag.*, pp. 48–59, May 2010.

[17] P. Robertson, M. Angermann, and B. Krach, "Simultaneous localization and mapping for pedestrians using only foot-mounted inertial sensors," in *Proc. 11th ACM Int. Conf. Ubiquitous Comput.*, 2009, pp. 93–96.

[18] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[19] P. Robertson, M. Angermann, and M. Khider, "Improving simultaneous localization and mapping for pedestrian navigation and automatic mapping of buildings by using online human-based feature labeling," in *Proc. IEEE/ION Position Location Navig. Symp.*, May 2010, pp. 365–374.

[20] P. Robertson, M. Garcia Puyol, and M. Angermann, "Collaborative pedestrian mapping of buildings using inertial sensors and FootSLAM," in *Proc. ION GNSS*, Sep. 2011, pp. 1366–1377.

[21] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. Part 1," in *Proc. IEEE Int. Conf. Commun.*, May 1993, vol. 2, pp. 1064–1070.

[22] K. P. Murphy, Y. Weiss, and M. I. Jordan, "Loopy belief propagation for approximate inference: An empirical study," in *Proc. Uncertainty AI*, 1999, pp. 467–475.

[23] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann, 1988.

[24] FootSLAM Videos. [Online]. Available: http://www.kn-s.dlr.de/indoornav

[25] M. Kasevich and S. Chu, "Atomic interferometry using stimulated Raman transitions," *Phys. Rev. Lett.*, vol. 67, pp. 181–184, Jul. 1991.

[26] B. Canuel, F. Leduc, D. Holleville, A. Gauguet, J. Fils, A. Virdis, A. Clairon, N. Dimarcq, C. J. Bordé, A. Landragin, and P. Bouyer, "Six-axis inertial sensor using cold-atom interferometry," *Phys. Rev. Lett.*, vol. 97, 010402, Jul. 2006.

ABOUT THE AUTHORS

Michael Angermann (Member, IEEE) received the Diplom-Ingenieur degree from the Technische Universität München (TUM), Munich, Germany, in 1998 and the Doktor-Ingenieur degree from the University of Ulm, Ulm, Germany, in 2004.

He is a Research Scientist at the Institute of Communications and Navigation, German Aerospace Center (DLR), Wessling, Germany. He teaches at TUM and the University of the Federal Armed Forces in Munich. He spent a research semester at the Computer Science and Artificial Intelligence Lab (CSAIL), Massachusetts Institute of Technology (MIT), Cambridge, in 2008. His research interests are in Bayesian methods with applications in communication, localization, mapping, and general context estimation of humans and multiagent robotic systems.

Patrick Robertson (Member, IEEE) received the Dipl.-Ing. degree in electrical engineering from the Technische Universität München (TUM), Munich, Germany, in 1989 and the Ph.D. from the University of the Federal Armed Forces, Munich, Germany, in 1995.

Since 1990, he has been working at the Institute for Communications and Navigation, German Aerospace Centre (DLR), Wessling, Germany. He has been active in the fields of mobile packet transmission, digital terrestrial television (DVB-T), multimedia transmission, synchronization, Turbo coding, and broadband wireless indoor communications networks. Since January 1999, he has been the leader of the Cooperative Systems research group. His current research interests include navigation signals and systems, simultaneous localization and mapping (SLAM), vehicle-to-vehicle communications networks and driver assistance services, and Bayesian inference techniques in these domains.