Event-Domain Knowledge in Inertial Sensor Based State Estimation of Human Motion

Tom Lucas Koller, Tim Laue and Udo Frese

March 28, 2022

Event-Domain Knowledge in Inertial Sensor Based State Estimation of Human Motion

1st Tom L. Koller
Mathematics and Computer Science
University of Bremen
Bremen, Germany
https://orcid.org/0000-0001-6629-1566

2nd Tim Laue
Mathematics and Computer Science
University of Bremen
Bremen, Germany
https://orcid.org/0000-0002-2845-1300

3rd Udo Frese
Mathematics and Computer Science
University of Bremen
Bremen, Germany
https://orcid.org/0000-0001-8325-6324

***Abstract*—State estimation can significantly benefit from prior knowledge about a system's dynamics and state. In this paper, we investigate a special class of prior knowledge: Events that correspond to a subset of the state space. This class of knowledge was introduced in pedestrian activity classification to improve the position estimation. We argue that the methodology can be generalized and applied to other applications in human motion tracking, in which the same class of knowledge is available. We apply this methodology to estimate the pose of climbers using inertial sensors and previously measured route maps. For our evaluation, we collected an open source dataset with 27 participants, including IMU data and ground truth positions of the hands. We detect gripping holds (as events), estimate the transition between holds in a least squares optimizer and use a particle filter to deploy the route map constraints (as state subset). In this scenario, our approach achieves a position median of $0.133\,m$ and thus demonstrates its possible effectiveness for this application class.**

***Index Terms*—IMU, INS, prior knowledge, PDR, bouldering, ZUPT, event-domain knowledge, human motion tracking**

I. Introduction

Model-based state estimation uses state, dynamic, sensor and noise models to estimate the probability distribution of the system's state. In many cases, prior knowledge is available, which is specific for an application and exceeds the modeling of sensor phenomena. This knowledge can greatly improve state estimation as it constrains the possible state space and thus acts as a prior on the probability distribution. It can be roughly categorized into knowledge that constrains the degrees of freedom (equality constraints) and knowledge that constrains the domain of the state (inequality constraints) [1]. Both types have been shown to be particularly useful to reduce the growing error of Inertial Navigation Systems (INS) even without additional sensors [2]–[4].

In pedestrian dead reckoning (PDR), the knowledge of the human motion pattern is essential to enable tracking with an inertial measurement unit (IMU) [4]. Whenever the foot is set on the ground, its velocity is 0 [5]. This knowledge enables to reset the velocity error and to estimate the IMU biases [6]. As a consequence, the drift of the position estimate is reduced. This zero velocity update (ZUPT) is conceptually different from knowledge like building maps [7], vehicle motion [2] or joint constraints [8] since it is only valid to be applied, when the foot is on the ground. Hence, it is coupled to an event.

DFG-Funding: ZaVI FR 2620/3-1

Fig. 1. Experimental evaluation in a bouldering scenario with IMUs and ground truth markers attached to the climbers' hands, feet and hip. The approach presented in this paper uses the hand IMUs only and estimates the positions of both hands separately. The estimates for the depicted situation are marked with green and red dots.

Location-based events such as walking stairs, riding elevators or opening doors are exploited in pedestrian activity classification (PAC) [9]–[11]. These actions are tied to structures or locations in the world. Thus, they constrain the possible state domain. Since multiple structures of the same type may exist, the constrained state domain may be multi-modal. In [11], they achieve a mean tracking error of $1.3\,m$ by matching detected turns to possible locations in the map.

In this paper, we want to investigate knowledge that is tied to events and constrains the state domain like in PAC. We call this kind *event-domain knowledge (EDK)*. EDK is closely related to landmark-based localization or SLAM where the landmark's position is the event domain [12]. The difference is that the event of being at a landmark is detected instead of computing a spatial relation from sensor measurements.

We want to show that EDK can be applied to achieve reasonable tracking performance in real-world scenarios other than PDR. We focus on human motion tracking in sports with IMUs as the only sensors. The motion in sports is restricted by rules and specific motions are used, wherefore EDK may be available. For instance, at hurdling, participants jump over the hurdles, which are at predefined positions [13]. Similarly, horses pass different obstacles in show jumping [14].

To achieve reasonable tracking performance, it is important that the transition between two event locations is well observed. If the estimation error of the transition is too high, it would not be possible to estimate at which event location the system is after a transition. Hence, applications get more challenging with longer time spans between the events.

Overall, indoor climbing is a fitting example for event-domain knowledge. The athletes climb up a wall by using the holds of a so-called route [15]. Hence, the motion can be split into periods of gripping holds and short periods of transitions. When gripping, the hands are constrained to be at a hold or a wall feature. Thus, we can infer information about the position while gripping. Furthermore, the velocity during a grip is zero, wherefore the ZUPT can be used to estimate the velocity. Since the transitions between holds are usually short, the event-domain knowledge can be applied regularly. Thus, it is likely that tracking is possible.

Indoor Bouldering (see Fig. 1) - a subcategory of climbing - is particularly suited for state estimation experiments. Since the routes are much shorter than in classic top rope or lead climbing, the sensor setup is significantly simplified. Furthermore, there is currently a high scientific interest in bouldering. Machine learning has been applied to classify different motion modes including gripping [16] and to detect which route is taken [17]. They try to create wearables that rate climbers' performance or analyse routes popularity [16], [17].

In this paper, we transfer the methodology of PDR and PAC to bouldering. We show that the event-domain knowledge is sufficient for tracking on real data. Furthermore, we identify challenging cases in the dataset. Since we see bouldering as an example for a whole class of applications, we want to emphasize that the retrieved methodology can be applied to other applications with EDK. Therefore, we will point out what we believe is the method's core that can be transferred between applications: To split the state estimation based on the observability.

In the next section, we will discuss how EDK can be applied in general based on the methodology in PDR and PAC. We use this methodology at bouldering in Section III and evaluate its effectiveness and challenges in Section IV. At the end, we summarize our findings and give future directions.

II. State Estimation with Event-Domain Knowledge

We start by characterizing systems with event-domain knowledge (EDK). Let $x_k \in X$ be the state of the considered system at time k. Consecutive states are connected by the system's dynamic model:

$$x_{k+1} = g_k(x_k, u_k) + \epsilon_k, \tag{1}$$

where u_k is the system's input and $\epsilon_k \sim \mathcal{N}(0, \Sigma_\epsilon)$ is the error of the dynamic model.

The states may be constrained by measurements h_k or a-priori distributions ϱ_k (prior knowledge). The state is further constrained by EDK. The EDK only applies to a subset of the state sequence because it is only valid if the corresponding event occurs. A selector function $d(k)$ with:

$$d(k) = \begin{cases} 1, & \text{if event occurs at time } k \\ 0, & \text{else} \end{cases}, \tag{2}$$

defines which states x_k are constrained by the EDK. This system is shown in Fig. 2.

The problem structure in PAC is an example of the presented EDK structure. In PAC, IMUs are used to detect steps and their direction, a so-called step and heading system (SHS) [9], [11]. The system detects turns, escalators and stairs in the IMU data. These events are matched to a map of possible event locations using a Hidden Markov Model (HMM) with the Viterbi algorithm. The approximated walking distance is used to compute the transition probability between event locations. Position, heading and step length are corrected with the matched event location.

Classic PDR methods use the wall map to constrain the position rather than the event locations. Although this only changes the structure from *event-domain* knowledge to *domain*-knowledge, the methodology is different in PDR. Sampling methods like the *particle filter (PF)* are widely used to handle the non-Gaussian probability distribution of the position [4]. The PF uses particles to represent the probability distribution. The number of required particles grows fastly with the number of dimensions of the state space [18].

The interesting common ground in PDR and PAC is to split the estimation. At first, they estimate the states of the transition without the map. Then, they correct all unobservable state like the position, heading and step length using the map. In both, the dimension is reduced since they fuse the map with position offsets rather than IMU measurements.

Inspired by PDR and PAC, we build our methodology upon the concept of splitting the estimation into a transition estimator like the step estimator and domain sampling like the PF. This is similar to Rao-Blackwellization, where parts of the state are represented with parametric distributions to reduce the sampled state dimension [19]. The drawback of splitting is that the PF does not propagate the information gain to the step estimator, e.g. the maps do not improve the IMU biases.

The information loss is negligible if the splitting is done correctly. In PDR, the velocity and inclination, as well as the IMU biases are observable due to the ZUPT [6]. Thus, the estimation error is already low and can only benefit slightly from the building maps. Therefore, we propose to split the estimation based on the observability of the states. By

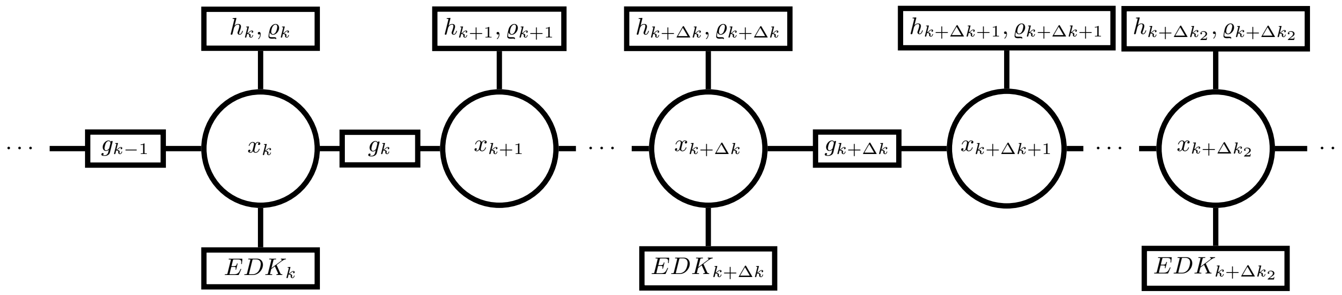

Fig. 2. The problem description as a factor graph. g_k is the dynamic model (1) given all inputs at time k, h_k are possible measurements, ϱ_k are possible a-priori distributions and EDK_k is event-domain knowledge that applies only to selected states, where $d(k) = 1$. Δk and Δk_2 are the index offsets where the EDK is applied.

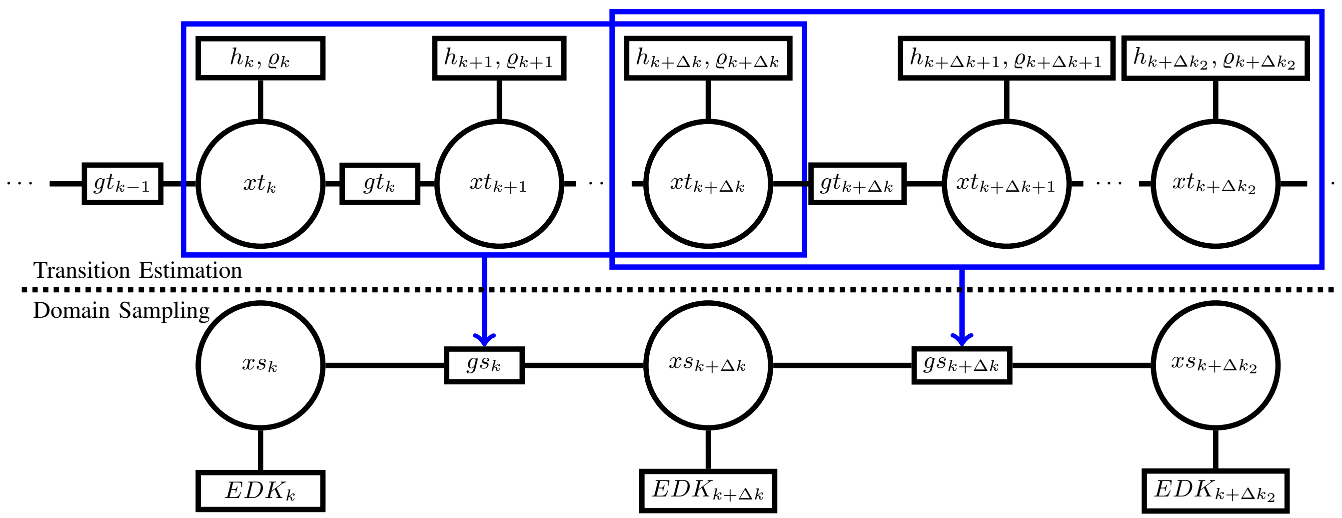

Fig. 3. The proposed methodology as a factor graph. gt_k is the dynamic transition model given all inputs at time k, gs_k is the dynamic sampling model which is given the result of the transition estimation (blue boxes), h_k are possible measurements, ϱ_k are possible a-priori distributions and EDK_k is event-domain knowledge that applies only to selected states. Δk and Δk_2 are the index offsets where the EDK is applied.

observable states we mean that the states are observable given all information except the sampled event-domain knowledge. These states may get observable by the EDK.

To sum it up, we propose to use the following methodology to apply event-domain knowledge (see Fig. 3):

1) *Detector Stage*: Detect the relevant events $d(k)$.
2) *Transition Estimation*: Estimate the observable transition states xt_k aided by measurements and a-priori distributions.
3) *Domain Sampling*: Sample the unobservable sampling states xs_k aided by the event-domain knowledge.

The detector stage is the basis to use EDK. It is important that the detector has as little false positive detections as possible. The EDK is only valid, if the event occurs. Thus, false information is introduced if the domain constraint is applied at other times.

The detection algorithm can be chosen freely for the given event. It is a perfect opportunity to use machine learning techniques in state estimation. Nevertheless, detectors based on statistics or custom metrics are widely spread and usable for event-domain knowledge [4], [9], [11].

As stated, the estimation task is split into transition estimation for the observable states and domain sampling for the unobservable states. Since the states in the transition estimation are observable, their distributions are usually unimodal and may be approximated by Gaussians. Hence, Kalman Filter variations, as well as least squares estimators are suitable choices for the transition estimation. Prior knowledge like the ZUPT or the non-holonomic constraint for wheeled vehicles can be easily applied as so-called pseudo measurements [20] to yield the necessary observability of the transition.

The essence of the domain sampling is that it can represent the multi-modal distribution of the event-domain. Hence, the domain sampling does not have to be a PF-style filter. Bayesian Occupancy grid (BOG) filters [21] are also possible.

III. State Estimation on Bouldering

As stated previously, bouldering is a prime example for event-domain knowledge and the pose of an IMU can be estimated using our methodology without using any sensors besides the IMU and the magnetometer. The athletes climb up a wall by using holds or the shape of the wall itself. Thus, their hands' movement is a sequence of gripping and transitions between holds. During the transitions, a hand's motion is unconstrained in 3D space. Thus, it is a standard INS System with the state space $X \subseteq SO3 \times \mathbb{R}^{15}$:

$$X = \begin{pmatrix} q_W^I & \vec{p}_W & \vec{v}_W & \vec{b_a} & \vec{b_g} & \vec{b_m}, \end{pmatrix}^T \tag{3}$$

where q_W^I is the orientation quaternion that rotates IMU to world coordinates, $\vec{p}_W$ is the position and $\vec{v}_W$ the velocity in 3D world space; and $\vec{b}_a$, $\vec{b}_g$, $\vec{b}_m$ are the biases of the accelerometer, gyrometer and magnetometer respectively. In this paper, we model the system's state as a so-called compound $\boxplus$-manifold (pronounced boxplus) [22]. This enables us to keep the quaternion as it is in the state without destroying its manifold properties during state estimation as long as we only use the accompanying $\boxplus$-operator ($\boxplus : X \times \mathbb{R}^{18} \mapsto X$) to apply changes to the state and to express the covariance in the tangent space. Concretely, this means that we express a change of the orientation as a rotation vector while representing the actual orientation with a quaternion. This prevents singularities in the state representation while taking into account that the orientation has only 3 DOF. The $\boxplus$-operator for the state X is:

$$x \in X, \delta \in \mathbb{R}^{18}: \quad x \boxplus \delta = \begin{pmatrix} q_W^I * \exp\left(\frac{\delta_{1:3}}{2}\right) \\ x_{vec} + \delta_{4:18} \end{pmatrix}, \tag{4}$$

where δ is an element of the tangent space of X, x_{vec} is the vector part of the state, $\exp(\cdots)$ is the exponential map of a rotation axis to a quaternion. The inverse operation is the boxminus operator ($\boxminus : X \times X \mapsto \mathbb{R}$):

$$a, b \in X,: \quad b \boxminus a = \begin{pmatrix} 2\overline{\log\left(a^{-1} * b\right)} \\ b_{vec} - a_{vec} \end{pmatrix}, \tag{5}$$

where $\log(\cdots)$ maps a rotation quaternion to a rotation vector and $\overline{\cdots}$ is the quaternion conjugation. This operator computes the tangential difference also called the geodesic. The dynamic model $g : X \times \mathbb{R}^3 \times \mathbb{R}^3 \mapsto X$ is:

$$g(x, \vec{a}_I, \vec{\omega}_I) = \begin{pmatrix} q_W^I * \exp\left(\frac{(\vec{\omega}_I - b_g)\cdot \Delta T}{2}\right) \\ \vec{p}_W + \Delta T \vec{v}_W + 0.5 \Delta T^2 \vec{a}_W \\ \vec{v}_W + \Delta T \vec{a}_W \\ \vec{b}_a \\ \vec{b}_g \\ \vec{b}_m \end{pmatrix} \boxplus \epsilon_g \tag{6}$$

$$a_W = q_W^I * (\vec{a}_I - \vec{b}_a) + \vec{g}, \tag{7}$$

where ΔT is the time difference between two IMU measurements, $\vec{a}_I$ is the acceleration and $\vec{\omega}_I$ the angular rate measured by the IMU, $\vec{g}$ is the gravity vector and $\epsilon_g \sim \mathcal{N}(0, \Sigma_\delta)$ with $\Sigma_\delta \in R^{18\times 18}$ as the dynamic covariance of X. Note that we use $*$ to denote standard quaternion multiplication between two quaternions but as rotation multiplication ($q * (0\ v) * \overline{q}$) when the right operand is a vector. Here, we model the IMU biases as constant but uncalibrated, as each trial of bouldering is rather short.

The measurement model $m : X \mapsto \mathbb{R}^3$ of the IMU magnetometer is:

$$m(x,t) = \overline{q}_W^I * (m_W + m_\epsilon(\vec{p}, t)) + \vec{b}_m + \mathcal{N}(0, \Sigma_m), \tag{8}$$

where $\overline{q}_W^I$ is the conjugation of q_W^I, m_W is the magnetic field vector in world coordinates and $\Sigma_m \in \mathbb{R}^{3\times 3}$ is the covariance of the magnetometer measurement. The magnetic field vector is usually disturbed by external magnetic fields m_ϵ which can vary over time and for each position. Without additional measurements, this INS system diverges fastly, since the gravity direction is unobservable which causes an enormous velocity error.

When gripping, the hands are constrained: Their velocity is close to zero and they have to be at a position where gripping is possible, such as a hold or a wall segment. Having zero velocity allows to correct the IMU biases and to estimate the direction of gravity. It is equivalent to the ZUPT in PDR, modeled by:

$$\vec{v} = \vec{0} + \mathcal{N}(0, \Sigma_{zupt}), \tag{9}$$

where Σ_{zupt} is the covariance of the assumption. Thus, the gravity direction, accelerometer and gyrometer biases, and the transition length can be estimated [6].

Consequently, the unobservable part of the state is the position. Due to the small errors of the velocity estimation, the position error would grow over time. We correct this error growth with the event-domain knowledge that athletes grip at hold positions. We follow our methodology presented in Section II to utilize the available knowledge:

1) We detect the event of gripping based on the IMU data using the method of [16]. This detector detects low acceleration phases in a moving window approach.
2) We use the INS model aided by the ZUPT in a transition estimator to estimate the velocity and sample the position distribution in a PF.
3) We sample not only the position but also the yaw error in the PF since the magnetic field disturbances may be unobservable and can cause a significant yaw error. We use an interval smoothing approach so that the information from later grips is backpropagated.

A. Transition Estimation

We chose a least squares approach to estimate the transitions. Least squares solvers automatically backpropagate information, wherefore the transition estimates between grips are optimized. Furthermore, their iterative nature improves the results in nonlinear estimation tasks.

We model the problem as a series of states x_k for each time point k which are connected by the dynamic model $gt_k = g(x_k, \vec{a}_{Ik}, \vec{\omega}_{Ik})$. Each state x_k is constrained by the magnetometer measurement m_k but only the states at which a grip is detected are constrained by the ZUPT. We cannot observe the magnetic disturbances, wherefore we use a simplified version of (8):

$$m(x) = \overline{q}_W^I * (m_W) + b_m + \mathcal{N}(0, \Sigma_m). \tag{10}$$

Since the magnetometer bias b_m may be unobservable we add an additional prior to prevent unrealistic bias estimates:

$$b_m = \mathcal{N}(\vec{0}, \Sigma_{bm}), \Sigma_{bm} \in \mathbb{R}^{3\times 3} \tag{11}$$

We estimate the optimal state series $\hat{x}$ with the Ceres solver [23] using the cost function:

$$\hat{x} = \arg\min_x \sum_{k=1}^{n} \Bigg[\underbrace{||g(x_k, \vec{a}_{Ik}, \vec{\omega}_{Ik}) \boxminus x_{k+1}||^2_{\Sigma_{\delta k}}}_{\text{dynamic model}} + \underbrace{||m(x_k) - m_k||^2_{\Sigma_m}}_{\text{magnetometer}} + \underbrace{d(k)||\vec{v}_{Wk}||^2_{\Sigma_{zupt}}}_{\text{ZUPT}} \Bigg] + \underbrace{||b_m||^2_{\Sigma_{bm}}}_{\text{magnet bias prior}} . \tag{12}$$

We did not split up the state sequence at the detected grips because this is not required for an offline estimation. Hence, the transition estimates benefit from the information gain of the whole trial to estimate the sensor biases.

Furthermore, we approximate the covariance of the transition $\Sigma_{\hat{x}}$ using the built-in tools of Ceres [23]:

$$\Sigma_{\hat{x}} = \left(J^T(\hat{x}) S_k^{-1} J(\hat{x})\right)^{-1} \tag{13}$$

$$S_k = diag\left(\Sigma_{\delta k} \quad \Sigma_m \quad \Sigma_{zupt} \quad \Sigma_{bm}\right) \tag{14}$$

$$J(x) = \frac{d}{dx} \sum_{k=1}^{n} \begin{pmatrix} g(x_k, \vec{a}_{Ik}, \vec{\omega}_{Ik}) \boxminus x_{k+1} \\ m(x_k) \\ \vec{v}_{Wk} \\ b_m \end{pmatrix} \tag{15}$$

In addition to the random noise of the accelerometer and gyrometer we add their nonlinearity error (percentage of sensor value) to the covariance since abrupt movements are common in bouldering. In the case of sensor range violations, e.g. excessive accelerations, we add a high covariance on the measurement to reflect the poorly defined measurement model outside the specification. The values of all covariances can be found in the Appendix.

In practice, least squares solvers can suffer from bad initialization, in particular if the cost function is nonlinear. Thus, we use the following procedure to acquire a suitable initial guess:

1) We precompute the orientations and magnetometer bias using the orientation estimator of [24].
2) We set the velocity to 0 at every grip.
3) We pass the states through the dynamic model (6) to compute the remaining states.

The transition estimator outputs the transition velocity $\vec{v}^*_W$ at every time step instead of the position displacement of the whole transition as the output of the transition estimator so that we can estimate the state between grip events. Due to unknown magnetic disturbances this velocity has a yaw error compared to the route map.

B. Domain Sampling

As stated, we sample only the position and the yaw error $\Delta\psi$ in the PF. Hence, given a set of M particles, each particle $x^{[m]} \in \mathbb{R}^4$ is defined as

$$x^{[m]} = \begin{pmatrix} \vec{p}^{[m]}_W & \Delta\psi^{[m]} \end{pmatrix}^T . \tag{16}$$

The dynamic model $g_P : \mathbb{R}^4 \times \mathbb{R}^3 \mapsto \mathbb{R}^4$ rotates the transition velocity $\vec{v}^*_W$ of the transition estimator by the yaw error:

$$g_P(x^{[m]}, \vec{v}^*_W) = \begin{pmatrix} \vec{p}^{[m]}_W + \Delta T \exp\begin{pmatrix} 0 \\ 0 \\ \Delta\psi^{[m]} \end{pmatrix} * \vec{v}^*_W \\ \Delta\psi^{[m]} \end{pmatrix} + \epsilon_p, \tag{17}$$

where $\epsilon_p \sim \mathcal{N}(0, \Sigma_P)$ with $\Sigma_P \in \mathbb{R}^{4\times4}$ consisting of the velocity covariance of the transition estimator and a covariance for the yaw error.

We use a maximum logic for the likelihood function of the grip so that only the likelihood of the most likely position is used. Thus, each particle implicitly represents the likelihood of state x given a sequence of grip positions rather than the cumulated likelihood of all possible sequences. This roughly corresponds to the HMM and Viterbi usage in [11] which computes the most likely sequence.

It is likely that the hand is at a hold of the taken route when gripping. However, in some routes, climbers grip at the wall to stabilize themselves or use graspable features of the wall. To account for this ambiguity, we model the likelihood distribution of gripping as the maximum of the likelihood of gripping a hold $h(x^{[m]})$ and gripping elsewhere $e(x^{[m]})$:

$$p(X = x^{[m]} | d(x^{[m]}) = 1) = \max(h(x^{[m]}), e(x^{[m]})). \tag{18}$$

Since we do not have a map of all possible grip position beside the holds, we use a uniform distribution for $e(x^{[m]})$:

$$e(x^{[m]}) = c_e, \tag{19}$$

where c_e is a constant likelihood.

The hold positions in the map are the center positions $\vec{p}_i$ of the holds. Often, the holds are not gripped at the center. Therefore, we use a ball shaped uniform distribution around each hold to avoid pulling the particles towards the center. Due to the variety in gripping techniques, it is not possible to safely distinct gripping a hold or elsewhere from the position alone. We added Gaussian tails to the uniform distribution to account for this, technically resulting in the convolution of a uniform distribution with a Gaussian around each hold:

$$h_i(x^{[m]}) = \begin{cases} \mathcal{N}\left(\vec{0}; \vec{0}, \Sigma_h\right), & \text{if } |\delta| < r \\ \mathcal{N}\left(\left(1 - \frac{r}{|\delta|}\right)\delta; \vec{0}, \Sigma_h\right), & \text{else} \end{cases} \tag{20}$$

$$\delta = \vec{p}^{[m]}_W - \vec{p}_i \tag{21}$$

where Σ_h is the covariance of the Gaussian tails and r is the radius of the uniform distribution. Again, we use the maximum to form the likelihood function for all holds:

$$h(x^{[m]}) = \max_i h_i(x^{[m]}) \tag{22}$$

We expect the probability distribution to be multi-modal. Therefore, we choose the mean of the most likely cluster of particles instead of all particles as the output. To cluster the particles, we store the last grip position (hold index or elsewhere) of every particle. Hence, each cluster represents the particles that where at the same grip position at the last grip event. We sum up the weights of each cluster to compute the cluster probability and to find the most likely cluster.

We use a standard PF and PF smoother following [18] using the systematic sampling approach for resampling. The size M of the particle set is 200 and resampling is performed whenever the effective sample size (ESS) [18] is below 50% of M.

Our complete implementation can be found at https://github.com/TomLKoller/BaVI-pose-tracking[1].

IV. Evaluation

To evaluate the methodology, we collected sensor data in a boulder study with 27 participants (M:21 F:6 Age:26.1±6.8). The study took place at a commercial boulder venue. The participants were equipped with wireless IMU sensors of the Xsens Awinda System [25] on the backs of the hands, their feet and their hip with hook and loop fastener belts (see Fig. 1). As ground truth, a marker-based infrared position tracking system with millimeter precision has been deployed (ART TrackPack with DTrack2, 2 cameras, 2011). The infrared markers were mounted on the IMUs. The IMU and ground truth measurements were hardware-synchronized. Additionally, a synchronized video of each run was recorded[2]. The ground truth system suffers from occlusion of the markers by the participants. Thus, the ground truth is sparse.

We obtained a map of the route holds by measuring the positions with the ground truth system.

Since participants were involved, the study has been approved by the ethics committee of the University of Bremen. All participants gave written consent to partake in this study. They were explained the purpose of the study and the overall setup. The participants were told to follow the standard rules of bouldering and – regarding the sensor setup – to boulder as usual, to continue even if they rip off an IMU and to ignore the ground truth system.

Overall, the dataset contains 12 routes, including a special route which consists of holds of other routes to force more dynamic bouldering moves. 775 valid trials of the left hand and 769 of the right hand have been recorded.

We initialize the PF around both start holds of a route. We labeled the start time as the kickoff time of the second feet from the ground for each trial and start the estimation at the first detected grip after the start. Furthermore, we labeled the end of the ascent (reaching the top hold) and use the route map for the ascent and all route maps of a wall for the descent, since it is allowed to use all holds then. We do so to ensure that the applied knowledge is valid.

A. Overall Pose Estimation Results

The grip detector [16] requires tuned parameters. Thus, we selected the data of 4 participants to test and tune it. Only the data of the remaining 23 participants is used for the evaluation.

The results of the PF and the PF smoother can be seen in TABLE I. The medians of the PF and the PF smoother are lower in comparison to the transition-only estimator. Thus, the event-domain knowledge overall improves the tracking. The smoother improves the result again as it can backpropagate the information from the hold map.

[1]Upload after acceptance

[2]The dataset will be published open access

TABLE I
Results of the Transition-Only, PF and PF smoother

Algorithm	25% Qtl.	Median	75% Qtl.	RMSE	Max
Transition-Only	0.125	0.266	0.542	0.571	13.013
PF	0.108	0.145	0.320	0.444	4.170
PF smoother	0.101	0.133	0.233	0.413	4.170

The maximum error and RMSE are exceptionally high compared to the median. This indicates that the tracking fails in some circumstances. In Fig. 4, the estimation results are shown exemplary for one trial. The estimation is overall close to the ground truth, which shows that the grip position sequence has been estimated correctly. Thus, the position tracking is successful.

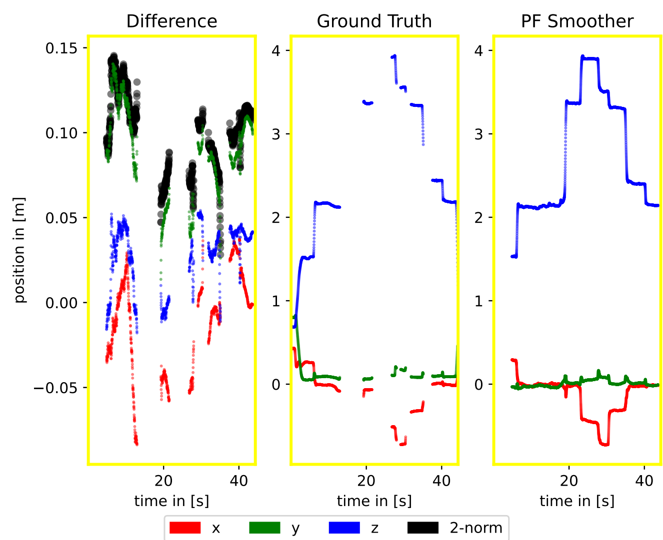

Fig. 4. Right hand position estimation of participant P53 in trial T30. The ground truth is shown at the center and the PF smoother result at the right. The left plot shows the error of the estimation.

Looking at all trials of the participant (see Fig. 5), this is not always the case. For example, T09 and T10 are exceptionally worse. These trials even have high minimum errors. Thus, the entire tracking failed in these trials. Furthermore, there are trials like the T23 or the T33, where the high 75% quantile indicates that the estimator lost track at some point. In the following, we will analyze why the tracking fails to give insights into the challenges with EDK.

B. Analysis of Challenging Cases

The transition estimation is impacted by outlier measurements. When participants stop abruptly at a hold, the measurement range of the IMUs is exceeded. The outliers violate the sensor model which is reflected in an overall worse performance. Trials with outliers have a higher overall RMSE (0.769) than trials without outliers (0.398) (see TABLE II).

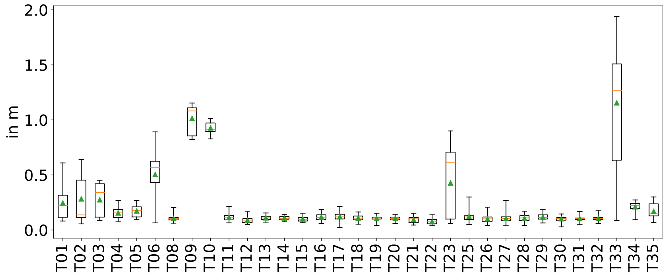

Fig. 5. Error distributions of the PF smoother results for participant P53's right hand on different routes.

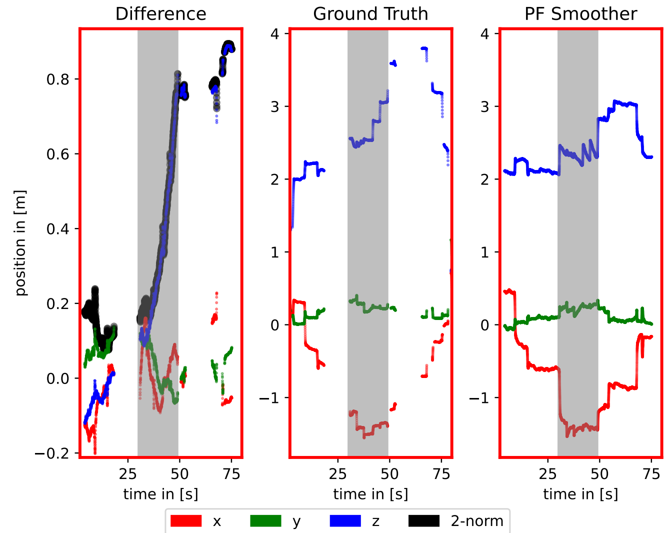

Fig. 7. Estimation of the right hand of P04 in trial T28. Grips at the wall between 30 s and 50 s (grey area). The position estimate is pulled wrongly to a close-by handle.

Similarly, local magnetic disturbances affect the heading estimate of the transition estimator.

TABLE II
METRICS OF THE PF SMOOTHER ON DIFFERENT TRIAL GROUPS.

	25% Qtl.	Median	75% Qtl.	RMSE	Max
With outliers	0.115	0.217	0.659	0.751	4.170
No outliers	0.101	0.131	0.216	0.382	2.961
Aborted	0.115	0.164	0.579	0.497	1.314
Completed	0.101	0.133	0.230	0.412	4.170

Aborted trials, in which the participant did not reach the top, have higher estimation errors than completed trials (see TABLE II). Aborted trials are shorter than completed, wherefore less transitions are taken. Thus, the start handle is often ambiguous due to a lack of information.

We assumed that all participants start the trial at the two starting holds. Hence, the PF can only localize the hand if it started there, which is wrong in some trials of the dataset. Furthermore, the PF suffers from particle depletion if the participants perform many grip adjustments at the start. Grip adjustments are not always detected as grips, wherefore the particles spread.

The error distribution appears to vary depending on the route (see Fig. 6). The *special* route has been designed to force dynamic bouldering moves. The transitions between grips are exceptionally long. Thus, the yaw error has a high impact on the estimation quality, wherefore the errors are overall high.

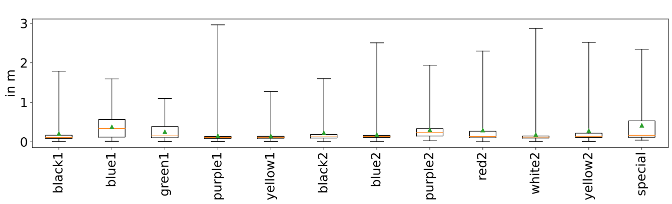

Fig. 6. Error metrics on the different routes without aborted or outlier trials. Routes with the same number are located on the same wall. Routes *green1* and *purple1* are almost identical as well as *blue1* and *yellow1*.

Sometimes, participants grip at wall features or simply place their hand on the wall to balance themselves. These grips improve only the velocity estimate but not the position. If these grips occur close to holds, a position bias is introduced into the system. The particles are pulled back to the hold and the tracking is lost (see Fig. 7). In particular the routes *purple2* and *red2* are affected by this, because wall grips are required to finish these. This behavior also shows, how false event detections increase the error where the estimate is wrongly pulled to a close-by handle.

At route *blue1*, the particles are often closer to the wrong handle, wherefore the tracking fails. Interestingly, this does not happen at route *yellow1* which is almost identical. We could not identify a specific reason of this regular failure.

The route *green1* consists of 2 almost vertical lines of holds. Here, the PF smoother is often unable to determine the correct starting hold which introduces an offset in the horizontal plane. Again, it is interesting that this does occur less at *purple1*, although it is almost equally shaped.

V. CONCLUSION

In this work, we introduced general systems subject to event-domain knowledge. We extracted the methodology for event-domain knowledge from PDR and applied it to bouldering. The core of the methodology is to split the estimation based on the observability, i.e. to estimate all observable states in a first step and sample all unobservable states in a second.

We presented bouldering as an application example for our methodology on EDK. We collected a real-world dataset with over 1500 single hand trials of IMU data with sparse position ground truth. The median error on the dataset is 0.133 m which shows the capability of EDK for tracking.

We identified different error sources which cause the tracking to deviate from the ground truth. Outliers of the IMU measurements impact the performance because they violate the noise models. Furthermore, violations of the knowledge-based assumptions falsify the probability distributions. In some trials, the start handle is ambiguous, due to a lack of information in the event-domain knowledge.

Overall, the presented methodology is a promising basis for the development of application specific algorithms. We combined standard state estimation techniques, namely least squares and PF, with simple models of the system and prior knowledge and achieved reasonable tracking performance on a challenging real-world dataset.

Our results showcase the overall capability of the method to track the position. Since we emphasized the general method, there are options to improve the results like a better modeling of the elsewhere distribution, of different hold sizes, a machine learning-based grip detector or sophisticated methods to handle erroneous sensor readings and magnetic disturbances. In future work, it can be tested whether it is possible to localize the hands with completely unknown start handles because this introduces more ambiguity into the system. Possibly, our methodology can be applied in smartphone-based PAC [11] by estimating position, yaw error and step length in the domain sampling instead of map matching via HMM.

APPENDIX

Covariance Matrices:

$$\Sigma_m = 10^{-2} I_{3\times3}, \quad \Sigma_{zupt} = 0.05^2 I_{3\times3}, \quad \Sigma_{bm} = 10^{-6} I_{3\times3}$$

$$\Sigma_{\delta k} = diag\left(\Sigma_{\delta qk} \quad 0_{3\times3} \quad \Sigma_{\delta vk} \quad 0_{9\times9}\right)$$

$$\Sigma_{\delta qk} = \left(\frac{\pi}{1800}\right)^2 I_{3\times3} + diag(10^{-3}|\omega_{Ik}|)^2 + c_k 100 I_{3\times3}$$

$$\Sigma_{\delta vk} = 4\cdot10^{-4} I_{3\times3} + diag(5\cdot10^{-3}|a_{Ik}|)^2 + c_k 900 I_{3\times3}$$

$$\Sigma_P = diag\left(\Sigma_{\hat{x}vk} \quad 0.09\right)\Delta T$$

where $|\cdots|$ is the coefficient wise absolute value and $c_k = 1$ if their is a sensor range violation at time k and else $c_k = 0$.

ACKNOWLEDGMENT

We thank the DAV-Kletterzentrum Bremen, its manager Jonas Loss and his stuff for providing the boulder gym and advise for the experiments and all participants of the boulder study. Furthermore, we want to honor the efforts of our students Lukas Post, Andreas Baude and Timo Wörner who set up the ground truth sensor system.

REFERENCES

[1] N. Amor, G. Rasool, and N. C. Bouaynaya, "Constrained State Estimation - A Review," arXiv:1807.03463 [eess], Dec. 2018, Accessed: Feb. 16, 2022. [Online]. Available: http://arxiv.org/abs/1807.03463

[2] 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 Transactions on Robotics and Automation, vol. 17, no. 5, Art. no. 5, Oct. 2001.

[3] T. L. Koller and U. Frese, "State Observability through Prior Knowledge: Analysis of the Height Map Prior for Track Cycling," Sensors, vol. 20, no. 9, Art. no. 9, Jan. 2020.

[4] R. Harle, "A Survey of Indoor Inertial Positioning Systems for Pedestrians," IEEE Communications Surveys Tutorials, vol. 15, no. 3, Art. no. 3, Third Quarter 2013.

[5] E. Foxlin, "Pedestrian tracking with shoe-mounted inertial sensors," IEEE Computer Graphics and Applications, vol. 25, no. 6, Art. no. 6, Nov. 2005.

[6] A. Ramanandan, A. Chen, and J. A. Farrell, "Inertial Navigation Aiding by Stationary Updates," IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 1, Art. no. 1, Mar. 2012.

[7] S. Beauregard, Widyawan, and M. Klepal, "Indoor PDR Performance Enhancement using Minimal Map Information and Particle Filters," in 2008 IEEE/ION Position, Location and Navigation Symposium, May 2008, pp. 141–147.

[8] F. Wenk and U. Frese, "Posture from motion," in 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Sep. 2015, pp. 280–285.

[9] S. Khalifa, M. Hassan, and A. Seneviratne, "Adaptive pedestrian activity classification for indoor dead reckoning systems," in International Conference on Indoor Positioning and Indoor Navigation, Oct. 2013, pp. 1–7.

[10] S. Grzonka, A. Karwath, F. Dijoux, and W. Burgard, "Activity-Based Estimation of Human Trajectories," IEEE Transactions on Robotics, vol. 28, no. 1, pp. 234–245, Feb. 2012.

[11] B. Zhou, Q. Li, Q. Mao, W. Tu, and X. Zhang, "Activity Sequence-Based Indoor Pedestrian Localization using Smartphones," IEEE Transactions on Human-Machine Systems, vol. 45, no. 5, Art. no. 5, 2015.

[12] F. Schuster, C. G. Keller, M. Rapp, M. Haueis, and C. Curio, "Landmark based radar SLAM using graph optimization," in 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Nov. 2016, pp. 2559–2564.

[13] B. McFarlane, "SPORTS PERFORMANCE SERIES: High performance hurdling–the women's 100m hurdles," Strength & Conditioning Journal, vol. 8, no. 6, pp. 4–16, Dec. 1986.

[14] H. M. Clayton, "Time-motion analysis of show jumping competitions," Journal of Equine Veterinary Science, vol. 16, no. 6, pp. 262–266, Jun. 1996.

[15] J. Boulanger, L. Seifert, R. Hérault, and J.-F. Coeurjolly, "Automatic Sensor-Based Detection and Classification of Climbing Activities," IEEE Sensors Journal, vol. 16, no. 3, Art. no. 3, Feb. 2016.

[16] C. Ladha, N. Y. Hammerla, P. Olivier, and T. Plötz, "ClimbAX: skill assessment for climbing enthusiasts," in Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, New York, NY, USA, Sep. 2013, pp. 235–244.

[17] F. Kosmalla, F. Daiber, and A. Krüger, "Climbsense: Automatic climbing route recognition using wrist-worn inertia measurement units," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, pp. 2033–2042.

[18] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," Handbook of nonlinear filtering, no. 12, pp. 656–704, 2009.

[19] S. Särkkä, A. Vehtari, and J. Lampinen, "Rao-Blackwellized particle filter for multiple target tracking," Information Fusion, vol. 8, no. 1, Art. no. 1, Jan. 2007.

[20] I. Klein, S. Filin and T. Toledo, "Psuedo-measurements as aiding to INS during GPS outages," NAVIGATION, Journal of the Institute of Navigation, vol. 57, no. 1, pp. 25-34, Spring 2010.

[21] M. Saval-Calvo, L. Medina-Valdés, J. M. Castillo-Secilla, S. Cuenca-Asensi, A. Martínez-Álvarez, and J. Villagrá, "A Review of the Bayesian Occupancy Filter," Sensors, vol. 17, no. 2, Art. no. 2, Feb. 2017.

[22] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," Information Fusion, vol. 14, no. 1, Art. no. 1, Jan. 2013.

[23] S. Agarwal, K. Mierle, and Others, "Ceres Solver." http://ceres-solver.org

[24] T. Seel and S. Ruppin, "Eliminating the Effect of Magnetic Disturbances on the Inclination Estimates of Inertial Sensors.," IFAC-PapersOnLine, vol. 50, no. 1, pp. 8798–8803, Jul. 2017.

[25] M. Paulich, M. Schepers, N. Rudigkeit, and G. Bellusci, "Xsens MTw Awinda: Miniature Wireless Inertial-Magnetic Motion Tracker for Highly Accurate 3D Kinematic Applications," p. 9.