Gravity aided navigation using Viterbi map matching algorithm

Abstract

In GNSS-denied environments, aiding a vehicle's inertial navigation system (INS) is crucial to reducing the accumulated navigation drift caused by sensor errors (e.g. bias and noise). One potential solution is to use measurements of gravity as an aiding source. The measurements are matched to a geo-referenced map of Earth's gravity in order to estimate the vehicle's position. In this paper, we propose a novel formulation of the map matching problem using a hidden Markov model (HMM). Specifically, we treat the spatial cells of the map as the hidden states of the HMM and present a Viterbi style algorithm to estimate the most likely sequence of states, i.e. most likely sequence of vehicle positions, that results in the sequence of observed gravity measurements. Using a realistic gravity map, we demonstrate the accuracy of our Viterbi map matching algorithm in a navigation scenario and illustrate its robustness compared to existing methods.

Keywords Gravity aided navigation · Gravity map matching · Hidden Markov Model · Maximum a posterior (MAP) estimation

1 Introduction

Since the advent of Global Navigation Satellite Systems (GNSS) in the 1960s, GNSS-based positioning has become ubiquitous in the areas of navigation, guidance and control. However, in a growing number of situations GNSS is either unavailable, e.g. in underwater environments [Wang et al. 2017a], or deliberately jammed or spoofed [Hemann et al. 2016]. In such GNSS-denied (or contested) situations, a platform's navigation performance is governed by on-board sensors – chief among them being the inertial navigation system (INS), which integrates inertial measurements from accelerometers and gyroscopes to obtain position. Unfortunately, even with high precision INSs, very small errors in the inertial sensors result in the build up of large navigation errors over time [Titterton & Weston 2004, Groves 2013]. Thus, an on-board aiding source that provides a position fix is crucial to removing these accumulated navigation errors and retaining confidence in the navigation accuracy. A promising candidate, and the focus of this paper, is the use of gravitational information to obtain a position fix [Han et al. 2016, Liu, Li & Jing 2019, Wang et al. 2022].
Gravity aided navigation originates from the early 1990’s [Affleck & Jircitano 1990, Jircitano & Dosch 1991]. The central concept is that an on-board sensor measures elements of the gravitational vector (or gravity gradient tensor) whilst the platform is in motion and these measurements are then matched to a geo-referenced map of the Earth’s gravitational field to determine a position. The advantage of such an approach is twofold: First, gravity is an inherent physical property of the Earth, thus it is immune to environmental interference, jamming and spoofing. Second, the measurement of gravity is a passive operation that does not require external information, which is advantageous in sensitive situations [Muniraj & Farhood 2017]. The challenge in this approach however is how to match the measurements to a position in the map. The design of this matching procedure, known as the map matching algorithm, is key to the success of gravity aided navigation.

The map matching algorithm needs to take into account: 1) the sensor noise when measuring gravity; 2) the uncertainty on the spatial location when a measurement is obtained; 3) possible ambiguity in the map resulting in a measurement being matched to multiple locations. In the literature, map matching algorithms tackle these problems in one of two ways. The first category of algorithms consider the problem as a single point matching operation. Building on the framework introduced in the Sandia terrain-aided navigation system (SITAN) [Hostetler & Andreas 1983, Bergman 1997], these techniques use the geo-referenced map as a look-up function and compute a predicted gravitational measurement using the INS estimated position. The predicted measurement is then used in the prediction step of a statistical filter. Due to the non-linear relationship between the estimation states and gravitational measurements, the extended Kalman filters (EKF) is often used to perform the estimation [Wei et al. 2017, Lee et al. 2015]. This category of algorithms take into account the sensor noise and spatial uncertainty but do not consider the map ambiguity directly. The second category of map matching algorithms directly considers the structure of the map. They approach the problem as a sequential matching operation. The classical example is the iterative closest contour point (ICCP) algorithm proposed in [Kamgar-Parsi & Kamgar-Parsi 1999] and utilised in [Wang et al. 2016, Han et al. 2017, Liu, Li, Lin & Jing 2019]. The ICCP is based on the observation that a single, scalar, gravity measurement creates an iso-contour of similar values in the gravity map space. Thus, given a sequence of gravity measurements, the matching problem is posed in terms of fitting a trajectory to a set of iso-contours based on initial position estimates and sensor measurements. To make the problem well posed, a regularization term is introduced to control the shape of the trajectory. However, linking the kinematic constraints of the platform’s motion to the regularization term is not straightforward.

In this paper, we propose a novel sequential map matching algorithm based on formulating the matching operation as a hidden Markov model (HMM). Specifically, we consider the spatial cells of the map to be the set of all possible hidden states in the model and the platform’s trajectory to be a particular sequence of these states. The observations from each state are the gravitational measurements and the output and state transition probabilities relate to the sensor noise and platform movement models, respectively. Thus, the map matching problem is equivalent to determining this sequence of states given a sequence of gravitational measurements. The advantage to this formulation is twofold: First, our proposed approach works directly in the map space and incorporates the modelling of the sensor noise and platform motion with the underlying structure of the gravity map. Second, this type of HMM problem can be solved efficiently using a dynamic programming method known as the Viterbi algorithm [Viterbi 1967, Ren & Karimi 2009]. Given the HMM formulation, the Viterbi algorithm estimates the most likely sequence of states that results in the sequence of observations and this sequence of states is optimal in the maximum a posterior (MAP) sense. Accordingly, we present a Viterbi map matching algorithm that outputs the most likely trajectory of the platform given a sequence of gravitational measurements. Furthermore, to tackle limited map resolution, we implement a two-layer, coarse-to-fine, estimation scheme to achieve sub-cell accuracy. We validate our algorithm using the ultra-high resolution, non-parametric, gravity maps presented in [Hirt et al. 2013] and demonstrate that it outperforms existing algorithms in terms of robustness and navigational accuracy. In particular, we demonstrate that our algorithm is able to tackle varying map spatial resolution as well as varying sensor noise levels.

2 Problem Formulation

Suppose that the actual position of a vehicle at time $k$ is $s_k = [x_k, y_k]^T$, the INS position estimate is $s_k^{INS} = [x_k^{INS}, y_k^{INS}]^T$ and the position corrected via map matching algorithm is $s_k^{C-INS} = [x_k^{C-INS}, y_k^{C-INS}]^T$. The gravity measurement at time $k$ and position $s_k$ is modeled as

$$z_k = g(s_k) + \epsilon_k$$  \hspace{1cm} (1)
where \( g(s_t) \) describes the true Earth's gravity at position \( s_t \) and \( \epsilon_t \sim N(0, \sigma_\epsilon^2) \) is noise with 0 mean and standard deviation \( \sigma_\epsilon \). Without loss of generality, in the simulation, the true Earth gravity is obtained from the gravity map without noise for the purpose of demonstrating the algorithm. Additionally, we assume that the INS provides velocity measurements \( v_k \) in the navigation frame and it is modeled as:

\[
v_k = v_k^0 + \phi_k
\]

where \( v_k^0 \) is the ground truth, \( \phi_k \) is the noise term comprising with a bias \( b_k \) and an independent zero mean Gaussian noise with co-variance \( \sigma^2 I_2 \) and \( I_2 \) is the \( 2 \times 2 \) identity matrix.

We assume that the INS position estimates are corrected after every segment of \( T \) gravity measurements. Without loss of generality, we assume that the INS position estimates and measurements are recorded from \( k+1 \). Then after \( T \) sampling times, i.e., at time \( k+T \), we have following available segments of measurements and INS position estimates:

\[
Z_k = \{z_{k+1}, \cdots, z_{k+T}\}, \quad \mathbf{V}_k = \{v_{k+1}, \cdots, v_{k+T}\}, \quad S_k^{INS} = \{s_{k+1}^{INS}, \cdots, s_{k+T}^{INS}\}
\]

The corrected position sequence using \((Z_k, S_k^{INS}, V_k)\) and the proposed map matching algorithm is denoted by \( \{c_{k+1}^{CINS}, \cdots, c_{k+T}^{CINS}\} \).

In this work, since the map matching algorithm is implemented every \( T \) time steps, therefore we actually are interested in how the algorithm works within time slot \([k+1, k+2, \cdots, k+T]\). For simplicity, where there is no confusion, time stamp \( k \) in this time slot is dropped from now on and we write \( t = 1, \cdots, T \).

### 3 Viterbi Map Matching Algorithm

#### 3.1 Constructing a grid on the gravity map

The gravity map can be viewed as a grid of pixels with each pixel representing an area equivalent to the resolution of the map. These pixels will henceforth be referred to as **cells**. The resolution of the gravity map, i.e. the dimensions of the cell, is \( R_x \times R_y \). Since each cell on the gravity map can be associated with a location, we simply let the central position of each cell be its location. Denote the collection of all cells' locations on the interested area of the map by \( C \). We define \( g_m() \) as the "look-up function" for finding the value from the gravity map using the given location, e.g. \( g_m(s_t^{INS}) \) is the gravity value associated to the cell that \( s_t^{INS} \) lies in.

In fact, at a specific sampling time \( t \), we are only interested in a limited area around the INS position estimate \( s_t^{INS} \) since the true position should be within that region. Therefore, given an estimate \( s_t^{INS} \), we can find a cell, or equivalently a pixel, on the geo-referenced map of gravity such that \( s_t^{INS} \) lies in. It should be noted that, since the cell is predefined and constructed according to the pixel of the map, and the estimate \( s_t^{INS} \) is arbitrary, therefore \( s_t^{INS} \) can be any point within the underlying cell and is unnecessary to be the center position of this cell. Then a grid of cells around the cell containing \( s_t^{INS} \) can be constructed and denoted by \( B_t \). Let the size of the grid be \( nR_x \times nR_y \), where \( n \) is a relatively small odd integer, and the positions of the cells within this grid be denoted by \( c_j^t \) for \( j = 1, \cdots, n^2 \). We also use \( c_j^t \) as the labels of the cells. The cells within the grid are arranged as shown in Fig.1 where the central cell of the grid, labelled as \( c_{\hat{n}}^t \) with \( \hat{n} = \frac{n^2+1}{2} \), is the one that \( s_t^{INS} \) lies in. Its position is

\[
c_{\hat{n}}^t \equiv [x_{\hat{n}}, y_{\hat{n}}]^T = \text{arg} \min_{[x,y]^T \in C} \| [x,y]^T - s_t^{INS} \|_2
\]

Essentially, \( c_{\hat{n}}^t \) means that \( c_{\hat{n}}^t \) is the cell that the INS position estimate \( s_t^{INS} \) lies in at time \( t \). As a result, each cell is associated to a gravity value according to the gravity map. Based on \( c_{\hat{n}}^t \) and \( n \), the positions of other cells in the grid \( B_k \) are

\[
c_j^t = c_{\hat{n}}^t - [m_1 R_x, m_2 R_y]^T, \quad \text{where} \quad m_1, m_2 = -n, \cdots, n \quad \text{and} \quad j = 1, \cdots, n^2
\]

#### 3.2 Viterbi based map-matching algorithm

At each time \( t = 1, 2, \cdots, T \), we construct the grid \( B_k \) and its associated cells \( c_j^t \), where the subscript \( t \) is again the time index and superscript \( j \) is the cell label \( j = 1, \cdots, n^2 \). Then the problem becomes how to select a cell
where \( I \) is the index set of the grid, \( c^i_t \) is the one that the INS position estimate \( s^i_{INS} \) lies in and can be determined via (3), and the positions of other cells can then be calculated by (4).

Note that, for arbitrary sequence, \( \{j_1^i, \cdots, j_T^i\} \), with corresponding referenced gravity values, \( \{g_m(c_1^i), \cdots, g_m(c_T^i)\} \), gravity measurements, \( Z_k \), INS information, \( V_k \) and \( S_k^{INS} \), the joint posterior probability is

\[
p \left( \{c_1^i, \cdots, c_T^i\} \mid Z_k, S_k^{INS}, V_k \right) = p \left( \{g_m(c_1^i), \cdots, g_m(c_T^i)\} \mid Z_k, S_k^{INS}, V_k \right) \tag{5}\]

The most probable trajectory from \( t = 1 \) to \( t = T \) is given by the MAP estimator

\[
j_1^i, \cdots, j_T^i = \arg \max_{T^x} p \left( \{c_1^i, \cdots, c_T^i\} \mid Z_k, S_k^{INS}, V_k \right) \tag{6}\]

where \( I = \{1, \cdots, n^2\} \), \( T^x = \times_{t=1}^T I \), notation \( \times \) is Cartesian product and \( \{j_1^i, \cdots, j_T^i\} \) is the optimal path (locations) of the vehicle estimated from the given measurements. It should be noted that \( c_1^i \) is actually the central position of the each cells of the map; that is, the estimated path for time interval 1, \( \cdots, T \) passes the center of each cell. Finally, the estimated position sequence is then used to reset the corresponding segment of this period estimated from INS.

The exhaustive searching process is required to find an optimal solution to (6), where the computation load of the order of \( n^T \) can explode quickly if \( n \) and/or \( T \) getting large. In this work, we intend to find an efficient algorithm to solve this problem.

Note that, for arbitrary sequence, \( \{j_1, \cdots, j_T\} \)

\[
p \left( c_1^j, c_{i-1}^j, \cdots, c_T^j \right) = p \left( c_1^j \mid c_{i-1}^j \right), \quad \forall t \in \{2, \cdots, T\} \text{ and } j_t \in I \tag{7}\]

which implies that the selection of \( c_1^j \) is only dependent on that of \( c_{i-1}^j \). This is a typical Markov process (Stroock 2013). The sequence \( \{c_1^j, \cdots, c_T^j\} \) with the measurements can be modelled by an HMM. The Viterbi algorithm (Viterbi 1967) is available to compute (6) and provides a most likely sequence of positions.

At each time \( t \), the likelihood function of \( z_t \) given \( c_t^j \) and \( g_m(c_t^j) \) is denoted by

\[
p_z \left( z_t \mid c_t^j \right) = p_z \left( z_t \mid g_m(c_t^j) \right) \tag{8}\]

where \( p_z(\cdot) \) is the probability density function of the gravity measurement. For each successive pair \( \{t-1, t\} \) with \( t = 2, \cdots, T \), the measurement likelihood function of \( c_t^j \) given \( v_{t-1} \), the velocity measurement at time \( t-1 \), and \( c_{t-1}^j \), is

\[
p_v \left( c_t^j \mid v_{t-1}, c_{t-1}^{j-1} \right) \tag{9}\]
where $p_v(\cdot)$ is the probability density function of the velocity measurement.

Then, given $c_{i-1}^{j-1}$ and measurement $z_t$, the posterior density of $c_{t}^{j}$ is

$$
p(c_{t}^{j}|z_{t}, v_{t-1}, c_{i-1}^{j-1}) \propto p(z_{t}|c_{t}^{j}, v_{t-1}, c_{i-1}^{j-1}) p(c_{t}^{j}|v_{t-1}, c_{i-1}^{j-1})$$

$$
= p_{x}(z_{t}|s_{m}(c_{t}^{j})) p_{v}(c_{t}^{j}|v_{t-1}, c_{i-1}^{j-1}) \quad \text{(10)}
$$

We shall be considering this, in particular, in the case when the transition is a single time step and $j_1 = j_{t-1}$ and $j_2 = j_t$, for a pair $(j_1, j_2) \in I \times I$, we use the notation

$$
\{j_1 \rightarrow j_2, p(c_{2}^{j_2}|z_{2}, v_{1}, c_{1}^{j_1})\} \quad \text{(11)}
$$

for the potential path from $c_{1}^{j_1}$ to $c_{2}^{j_2}$ and its associated conditional probability.

Similarly, for any pair $(j_{t-1}, j_t) \in I \times I$, we can have such a path with its posterior probability as

$$
\{j_{t-1} \rightarrow j_t, p(c_{t}^{j_t}|z_{t}, v_{t-1}, c_{t-1}^{j_{t-1}})\} \quad \text{(12)}
$$

Then, starting from $t = 2$, for a fixed $j_1$, a local optimal path with greatest posterior probability is given by

$$
\{j_1 \rightarrow j_2 | j_2 = \arg\max_{j_2 \in I} p(c_{2}^{j_2}|z_{2}, v_{1}, c_{1}^{j_1})\} \quad \text{(13)}
$$

Similarly, at $t = 3$, for pair $(j_2, j_3)$, where $j_2$ is obtained from [13] and therefore fixed and $j_3 \in I$, the following local optimal path is

$$
\{j_2 \rightarrow j_3 | j_3 = \arg\max_{j_3 \in I} p(c_{3}^{j_3}|z_{3}, v_{2}, c_{2}^{j_2})\} \quad \text{(14)}
$$

Iteratively, for any pair $(j_{t-1}, j_t), j_t \in I$, the following optimization is done:

$$
\{j_{t-1} \rightarrow j_t | j_t = \arg\max_{j_t \in I} p(c_{t}^{j_t}|z_{t}, v_{t-1}, c_{t-1}^{j_{t-1}})\} \quad \text{(15)}
$$

As a result, for any fixed $j_1 \in I$, there is a candidate path that the vehicle potentially travels:

$$
\{j_1 \rightarrow j_2 \rightarrow \cdots \rightarrow j_{t-1} \rightarrow j_t \} \doteq \mathcal{J}^{*}_{j_1} \quad \text{(16)}
$$

where

$$
j_t^* = \begin{cases} 
\arg\max_{j_t \in I} p(c_{t}^{j_t}|z_{t}, v_{t-1}, c_{t-1}^{j_{t-1}}) & t = 2 \\
\arg\max_{j_t \in I} p(c_{t}^{j_t}|z_{t}, v_{t-1}, c_{t-1}^{j_{t-1}}) & t = 3, \cdots, T 
\end{cases} \quad \text{(17)}
$$

Since $\mathcal{J}^{*}_{j_1}$ depends on the selection of $j_1$, we are able to collect the potential paths corresponding to all possible values of $j_1$, i.e. $\mathcal{J}^* = \{\mathcal{J}^{*}_{1}, \ldots, \mathcal{J}^{*}_{u}\}$. Then,

$$
j_1^* = \arg\max_{j_1 \in I} p_{z}(z_{1}|s_{m}(c_{1}^{j_1})) p_{v}(c_{2}^{j_2}|z_{2}, v_{1}, c_{1}^{j_1}) \prod_{t=3}^{T} p_{v}(c_{t}^{j_t}|z_{t}, v_{t-1}, c_{t-1}^{j_{t-1}}) \quad \text{(18)}
$$

and the complete estimated path is

$$
\{j_1^* \rightarrow j_2 \rightarrow \cdots \rightarrow j_{t-1} \rightarrow j_t \} \doteq \mathcal{J}^{*} \quad \text{(19)}
$$

Eventually, the estimated path $\mathcal{J}^{*}$ is used to reset the INS estimated position directly. An example of how the algorithm works is shown in Fig. 2.

In practice, given a measurement $z_t$ at a specific time $t$, we may find more than one cell within a grid $B_t$ that has the same likelihood value using [8]. Therefore multiple optimal trajectories may be identified because of that they can have the same maximum posterior value. If this happens we choose the trajectory that is closest to the INS estimated position for simplicity.
3.3 An enhanced two-layer algorithm

In Section 3.2, a grid-based Viterbi algorithm is proposed to find a path used to correct $S_{t}^{INS}$. However, it should be noted that the proposed algorithm only returns a path that passes through the center of each cell. As a result, the deviation between this path and the actual trajectory is highly dependent on the resolution of the gravity map since the actual trajectory can be any point of each cell. In this section, an enhanced method is presented to improve the accuracy of the estimated path.

The key idea of this enhanced algorithm is to generate refined sub-cells for each cell and use these refined sub-cells to match the actual trajectory. This strategy takes effect when constructing the grid $B_{t}$. An example of a refined cell is given in Fig. 3.

Figure 3: Illustration of the refined grid. The cell $c_{t}^{5}$ is an example to show how the refined sub-cells are constructed, where $n = 3$ and $o = 3$.

Suppose that each cell is divided into $o \times o$ sub-cells, the centers of which, for $c_{t}^{j}$, are denoted by $c_{t}^{j,l}$, where $l_{i} \in O$ and $O = \{1, 2, \ldots, o^{2}\}$. We also use $c_{t}^{j,l}$ as label of $l_{i}$-th sub-cell of $j_{i}$-th cell at time $t$. Since the
resolution of gravity map is fixed, therefore the sub-cells \( c_i^{j,l} \) refine cells \( c_i^j \) to provide better resolution on positions and do not improve the actual resolution of the gravity map. In the other word, the cell \( c_i^j \) within \( B_i \) is divided into multiple sub-cells \( c_i^{j,l} \). Then we have the following relation, for any given \( j_i \):
\[
g_m(c_i^j) = g_m(c_i^{j,l_i}), \forall l_i \in \mathcal{O}
\]
and
\[
p_z(z_i|g_m(c_i^{j,l_i})) = p_z(z_i|g_m(c_i^j)), \forall l_i \in \mathcal{O}
\]
In similar vein to the (15), (16) and (17), the potential path for a given \( \{j_i,l_i\} \) is
\[
\{j_1,l_1\} \rightarrow \{j_2,l_2\} \rightarrow \cdots \rightarrow \{j_{\tau-1},l_{\tau-1}\} \rightarrow \{j_{\tau},l_{\tau}\} \notin \mathcal{L} \quad (22)
\]
where \( l_i^* \) in \( \{j_i^*,l_i^*\} \) is the optimal sub-cell and \( j_i^* \) is the associated cell. \( \{j_i^*,l_i^*\} \) is obtained using \( \{j_{\tau-1}^*,l_{\tau-1}^*\} \) and it can be calculated from (17) via replacing \( j_i \) by \( j_i \).
Let \( \mathcal{L}^- = \{\mathcal{L}_{1,1}^-, \mathcal{L}_{1,2}^-, \cdots, \mathcal{L}_{n_x,z}^-\} \). Then \( \{j_i^*,l_i^*\} \) in the estimated path
\[
\mathcal{L}^* = \{j_{\tau}^*,l_{\tau}^*\} \rightarrow \{j_{\tau-1}^*,l_{\tau-1}^*\} \rightarrow \cdots \rightarrow \{j_2^*,l_2^*\} \rightarrow \{j_1^*,l_1^*\}
\]
is obtained by
\[
\{j_i^*,l_i^*\} = \arg \max_{\mathcal{L}^-} p_z(z_i|g_m(c_i^{j,i,l_i})) \left( c_{z_2}^{j_2,l_2^*} | z_2, v_1, c_{i_1}^{j_1,l_i} \right) \prod_{t=3}^{T} p\left( c_{z_2}^{j_2,l_2^*} | z_t, v_{t-1}, c_{i_{t-1}}^{j_{t-1},l_{t-1}^*} \right) \quad (24)
\]

### 3.4 Reducing the Computation

In the algorithms presented above, the computational overhead increases dramatically as more cells are considered in the algorithm. To mitigate this issue, we constrain the number of \( j_i \) in each time \( t \) via a threshold. Relative likelihood values to select the available \( j_i \) are calculated as follows:
\[
\mathcal{A}_t = \left\{ j_i \left| \frac{p_z(z_t|g_m(c_i^j))}{\max_{j_i \in \mathcal{T}} p_z(z_t|g_m(c_i^j))} \geq \alpha, \forall j_i \in \mathcal{T} \right. \right\} \quad (25)
\]
where \( \alpha \in [0,1] \). (25) implies that only those cells with relatively high likelihood values are chosen. In other words, cells that are very unlikely to form part of the path are discarded.

An example of higher likelihood cells from \( c_{i}^j \), \( j_i \in \mathcal{T} \), is given in Fig. 5, where we assume that \( T = 3 \) and the size of the grid at each time stamp is \( n = 11 \). The values of \( p_z(z_t|g_m(c_i^j)) \), \( t = 1,2,3 \), are represented in different colors. At time 1, there are three higher likelihood cells. Similarly, at time 2 and 3, there are 3 and 2 grids available respectively. As a result, \( \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3 \), is obtained accordingly, e.g., \( \mathcal{A}_1 = \{8, 18, 29, 40, 61, 83, 94\} \). The computational complexity is reduced from \( O((T-1)n^2) \) to \( O(\sum_{t=1}^T \# \mathcal{A}_t \times \# \mathcal{A}_{t+1}) \), where \( \# \mathcal{A}_t \) is the number of elements in \( \mathcal{A}_t \). In practice, to simplify the problem, \( \alpha \) can be set to be a fixed value, such as 0.1 or 0.2.

Similarly, we use indexes in \( \mathcal{A}_1, \cdots, \mathcal{A}_T \) to find multiple candidate paths using the same method as (22) by replacing the associated indexes sequence, i.e.,
\[
\{\{j_1^*,l_1^*\} \rightarrow \{j_2^*,l_2^*\} \rightarrow \cdots \rightarrow \{j_{\tau-1}^*,l_{\tau-1}^*\} \rightarrow \{j_{\tau}^*,l_{\tau}^*\} \notin \mathcal{L}^* \quad (26)
\]
where \( j_i \in \mathcal{A}_1 \) and \( j_i \in \mathcal{A}_t \) for \( t = 2, \cdots, T \).

Let \( \mathcal{R}^- = \{\mathcal{R}_{1,1}^-, \cdots, \mathcal{R}_{n_x,z}^-\} \) contain these candidate paths. Then (24) can be rewritten as, where \( j_i \in \mathcal{A}_1 \),
\[
\{j_i^*,l_i^*\} = \arg \max_{\mathcal{R}^-} p_z(z_i|g_m(c_i^{j,i,l_i})) \left( c_{z_2}^{j_2,l_2^*} | z_2, v_1, c_{i_1}^{j_1,l_i} \right) \prod_{t=3}^{T} p\left( c_{z_2}^{j_2,l_2^*} | z_t, v_{t-1}, c_{i_{t-1}}^{j_{t-1},l_{t-1}^*} \right) \quad (27)
\]
The estimated path is, then, as follows:
\[
\mathcal{R}^* = \{\{j_2^*,l_2^*\} \rightarrow \{j_1^*,l_1^*\} \rightarrow \{j_2^*,l_2^*\} \rightarrow \cdots \rightarrow \{j_{\tau-1}^*,l_{\tau-1}^*\} \rightarrow \{j_{\tau}^*,l_{\tau}^*\} \}, \text{ where } j_i \in \mathcal{A}_t \quad (28)
\]
In summary, the algorithm proposed in Sections 3.2 with enhanced strategies presented in 3.3 and 3.4 is given in Algorithm 1.
Navigation using Viterbi Map Matching

Figure 4: Illustration of likelihood values of cells within the grids at successive time stamp \( t = 1, 2, 3 \). The windows size is \( n = 11 \times 11 \). For arbitrary cell \( j_t \) at time \( t \), the likelihood values are calculated from 

\[
p_z(z_t|g_m(c^j_t)) \text{ and normalized via } p_z(z_t|g_m(c^j_t))/\max_{j'_t \in T} \{ p_z(z_t|g_m(c^{j'_t}) \}. \]

There is no sub-cell for each cell to simplify the illustration. In each grid, there are only up to 7 cells having relative high likelihood values while other cells' values are close to 0. If (25) is applied and let \( \alpha = 0.1 \), the maximum number of possible trajectories is \( 7^3 \); otherwise, in the algorithm without applying (25), the maximum number of possible trajectories is \( 121^3 \).

Algorithm 1: The Viterbi-based map matching algorithm

Result: The estimated trajectory

1. Initialization \( k \leftarrow 1 \)
2. while \( k < \text{Running time} \) do
3.   Obtain position via INS
4.   Push measurements and INS estimated position into vector \( Z_k, S^{\text{INS}}_k \) and \( V_k \) respectively.
5.   if \( k \mod T = 0 \) then
6.     Construct grids \( B_t \) using \( n, o_t \) \( (3) \) and sub-cell method proposed in Section 3.3 for each \( t = 1, \cdots, T \)
7.     Find all cells \( A_t \) satisfying (25) for each \( t = 1, \cdots, T \)
8.     for \( j_t \in A_t \) do
9.       for \( l_t \leftarrow 1 \) to \( o \) do
10.      for \( t \leftarrow 1 \) to \( T \) do
11.         Find ordered sequence \( R^*_j \) according to (26)
12.       end
13.       Push \( R^*_j \) into \( R^* \)
14.     end
15.     end
16.     Find the estimated path \( R^* \) using \( R^* \) via (28)
17.     Replace the INS estimated trajectory using \( R^* \)
18.     Clear \( Z_k, S^{\text{INS}}_k \) and \( V_k \)
19. end
20. \( k \leftarrow k + 1 \)
21. end
4 Simulation

In the first part of the simulation, we assume that a vehicle (aircraft) travels across Australia, approximately from Melbourne to Sydney, at a constant velocity $7.54^\circ /h$, roughly 837 Km/h ($1^\circ \approx 111$Km (Scales et al. 2017)). The total flight distance is about 814 Km. The gravity map of the region of interest has been obtained from the GGMplus software (Hirt et al. 2013) with a variety of resolutions, where the component taken from the map is gravitational acceleration. The map is represented in grids that have specified resolutions with the grid equally spaced in terms of geodetic (GRS80) latitude and longitude. The map with lower resolution is created by down-sampling the higher one. An illustrative example is shown in Fig. 5a.

To simplify the simulation, the noise in the velocity measurement, i.e. $\phi_k$, from the IMU is modelled as $\phi_k = b_k + \mathcal{N}(0,\sigma_v^2 I_2)$, where $b_k$ is bias vector. The measured gravitational acceleration is assumed to be corrupted by additive Gaussian noise $\mathcal{N}(0,\sigma_z^2)$, where $\sigma_z$ is the standard deviation. As mentioned earlier, the true Earth gravity is obtained from gravity map without noise for the purpose of demonstrating the algorithm.

A Monte Carlo (MC) simulation of 500 runs is carried out for each scenario and simulation performance at time $k$ is measured by the root-mean-square error, defined as

$$\text{Error}_k = \frac{1}{nMC} \sum_{m=1}^{nMT} \text{Dist}(s_{k,m}, s^{C-INS}_{k,m})$$

(29)

where $nMC$ is the number of MC, $s_{k,m}$ and $s^{C-INS}_{k,m}$ are the position and updated position respectively of the target at time $k$ and $m$-th simulation and Dist(·, ·) is the Haversine distance (Km) between two points given their longitude and latitude (Gade 2010).

The algorithms are evaluated by different parameters. In the simulation, the bias of the velocity is set to be $1^\circ/hr$ and $\sigma_v = 9 \times 10^{-6}/s$, roughly 1 m/s. In the simulation, since we assume that the inertial sensors used in the vehicle INS exhibits white noise and bias, such that the INS computed position deviates from ground truth at the end of journey as shown in Fig. 5. The standard deviation of the gravitational acceleration measurement is $\sigma_z = 1.2$ mGal. The INS reports positions and velocities every 12 s.

Firstly, the performances of the Viterbi based map matching algorithm (VBMP) proposed in Section 3.2 and the computation reduced algorithm (RVBMP) proposed in Section 3.4 are compared in terms of the time
efficiency. In the simulation, standard deviation $\sigma_z = 1$ mGal, the map resolution is $1/100^\circ$, the correction rate, i.e. $T$, is set to be 6, window size is $n = 13$ and the sub-cell algorithm proposed in 3.3 is not applied. It also can be noticed that the value of $\alpha$ can potentially have impact on the performance of RVBMP; that is, the larger the $\alpha$ is, the more cells will be included in $A_t$ and more time will be required to find a solution. In considering this, the performance of RVBMP using different values of $\alpha$ are compared with VBMP. When $\alpha = 0$, set $A_t$ will contain all cells and the algorithm becomes VBMP. The simulation results are given in Table 1, where TCR stands for time consumption ratio between the time consumption using a specific $\alpha$ and a referenced time, i.e. $\frac{\text{Time}_{\alpha}}{\text{Time}_{\alpha=0.4}}$, and Time$_{\alpha=0.4}$ is chosen as the reference in Table 1. Mean and Std. Dev. are sample mean and sample standard deviation of Error$_k$, i.e.

$$
\text{Mean} = \frac{1}{L} \sum_{k=1}^{L} \text{Error}_k, \quad \text{Std. Dev.} = \sqrt{\frac{1}{L-1} \sum_{k=1}^{L} (\text{Error}_k - \text{Mean})}
$$

where $L$ is the total number of sampling points over the trajectory.

Table 1: Performance comparison between RVBMP and RVBMP using different values of $\alpha$. The scenario is shown in Fig. 5a.

| $\alpha$ (VBMP) | Mean (Km) | Std. Dev. (Km) | TCR |
|-----------------|-----------|----------------|-----|
| 0 (VBMP)        | 1.2802    | 0.5011         | 35.3817 |
| 0.05            | 1.2890    | 0.5101         | 2.5162 |
| 0.1             | 1.2786    | 0.5219         | 2.0119 |
| 0.2             | 1.5114    | 0.5470         | 1.5049 |
| 0.3             | 1.7097    | 0.5655         | 1.1986 |
| 0.4             | 2.3764    | 0.8590         | 1     |

It can be seen that the RVBMP performs better than VBMP in terms of time efficiency and they have the similar tracking performance when $\alpha$ is less than 0.1.

Secondly, the performance of RVBMP and the enhanced two-layer algorithm (RVBMP-2) in Section 3.3 are compared via standard deviation $\sigma_z = 1$ mGal and two map resolutions. The correction rate, i.e. $T$, is set to be 6, window size is $n = 13$ and the size of a sub-cell is $\sigma = 7$ for RVBMP-2. The definitions of Mean and Std. Dev. are same as those in Table 1 and they are calculated over those estimated results that are not divergent. The time consumption ratio (TCR) is given to compare the time efficiency of these two algorithms. The TCR in this table is defined as: $\frac{\text{Time}_{\text{RVBMP-2}}}{\text{Time}_{\text{RVBMP}}}$.

From Fig. 6 and Table 2 we can see that the RVBMP-2 outperforms RVBMP in both resolutions, which indicates that the sub-cell can improve the RVBMP. It should be noticed that the sub-cell strategy works better when the map resolution is relative low. This is because of that the high map resolution implies that high position resolution as well so that the sub-cell strategy has less influence on the performance.

Table 2: Performance comparison between RVBMP and RVBMP-2 using two different map resolutions. The scenario is shown in Fig. 5a.

| Map resolution = $1/100^\circ$ | Algorithm | Mean (Km) | Std. Dev. (Km) | TCR |
|---------------------------------|-----------|-----------|----------------|-----|
| RVBMP                           | 1.4358    | 0.5715    | 1              |
| RVBMP-2                         | 0.9190    | 0.4004    | 24.9           |

| Map resolution = $1/200^\circ$ | Algorithm | Mean (Km) | Std. Dev. (Km) | TCR |
|---------------------------------|-----------|-----------|----------------|-----|
| RVBMP                           | 0.5846    | 0.2639    | 1              |
| RVBMP-2                         | 0.5266    | 0.2538    | 28.5           |

Finally, to further demonstrate the efficiency of RVBMP-2, the performance of RVBMP-2 and ICCP (Kamgar-Parsi & Kamgar-Parsi 1999, Han et al. 2017) are compared using two different noise levels, i.e. $\sigma_z$. The map resolution is $1/200^\circ$, window size is $n = 13$ and the size of sub-cell is $\sigma = 5$. The simulation results are
Figure 6: Simulation results. The scenario is shown in Fig. 5a. The map resolution is 1/100°; (b) The map resolution is 1/200°.

From Table 3 and Fig. 7-8 we see that RVBMP-2 algorithm outperforms ICCP both in success rate and tracking performance.

In the second part of the simulation, a reference gravity map that covers a major area of Western Australia (WA), Australia, see Fig. 5b, is used to further demonstrate the performance of the proposed algorithm. The major challenge of this scenario is that the gravitational acceleration is relatively flat; that is, the variation in gravity across distinct pixels/cells is, on the whole, smaller than that in the regions in the earlier simulations. This flatness of the gravity map suggests that aiding with it will perform less well than in the previous results.

In the simulation, the start point of the vehicle is near Perth, and the end point is located at (−24.427111, 124.978610). The distance is about 1206 Km. The remaining parameters are similar to those in Table 3.
Figure 8: Simulation results of ICCP and proposed algorithm when the correction rate is 6. The scenario is shown in Fig. 5a [(a)] $\sigma_z = 1$ mGal; [(b)] $\sigma_z = 2$ mGal.

Table 3: Performance comparison between RVBMP-2 and ICCP using two different standard deviations of the noise, $\sigma_z$ and correction rate. The scenario is shown in Fig. 5a.

| Correction rate = 4 | $\sigma_z$ (mGal) | Algorithm | Success rate | Mean (Km)  | Std. Dev. (Km) |
|---------------------|-------------------|-----------|--------------|-------------|----------------|
|                     | 1                 | ICCP      | 92%          | 1.3424      | 0.6642         |
|                     |                   | RVBMP-2   | 100%         | 0.8300      | 0.3557         |
|                     | 2                 | ICCP      | 54%          | 2.0253      | 1.1920         |
|                     |                   | RVBMP-2   | 100%         | 1.4238      | 0.4146         |

| Correction rate = 6 | $\sigma_z$ (mGal) | Algorithm | Success rate | Mean (Km)  | Std. Dev. (Km) |
|---------------------|-------------------|-----------|--------------|-------------|----------------|
|                     | 1                 | ICCP      | 75%          | 0.8675      | 0.4297         |
|                     |                   | RVBMP-2   | 100%         | 0.5765      | 0.2722         |
|                     | 2                 | ICCP      | 46%          | 1.1533      | 0.5148         |
|                     |                   | RVBMP-2   | 100%         | 0.8478      | 0.3466         |

Table 4: Performance comparison between RVBMP-2 and ICCP using two different standard deviations of the noise, $\sigma_z$ and the correction rate. The scenario is shown in Fig. 5b.

| Correction rate = 4 | $\sigma_z$ (mGal) | Algorithm | Success rate | Mean (Km)  | Std. Dev. (Km) |
|---------------------|-------------------|-----------|--------------|-------------|----------------|
|                     | 1                 | ICCP      | 23%          | 1.5880      | 0.8637         |
|                     |                   | RVBMP-2   | 100%         | 1.4861      | 0.6183         |
|                     | 2                 | ICCP      | 2%           | 2.4581      | 1.0113         |
|                     |                   | RVBMP-2   | 100%         | 2.3573      | 0.8325         |

| Correction rate = 6 | $\sigma_z$ (mGal) | Algorithm | Success rate | Mean (Km)  | Std. Dev. (Km) |
|---------------------|-------------------|-----------|--------------|-------------|----------------|
|                     | 1                 | ICCP      | 22%          | 1.0967      | 0.4457         |
|                     |                   | RVBMP-2   | 100%         | 1.0679      | 0.4336         |
|                     | 2                 | ICCP      | 2%           | 1.6697      | 0.7551         |
|                     |                   | RVBMP-2   | 99%          | 1.6003      | 0.5975         |

In the Table we can see that, in terms of accuracy, the performances of RVBMP-2 and ICCP are similar because of the map is relatively flat and provides less information to the algorithm. However, the former has better a success rate than the latter.
5 Summary

In this paper, we have presented a Viterbi-inspired map matching algorithm based on a hidden Markov model (HMM) formalisation of the map matching problem. Our algorithm determines the optimal trajectory of a platform given a sequence of gravitational measurements, a geo-reference gravity map and INS velocity information. This trajectory is optimal in the maximum a posterior sense and is used to correct the platform’s INS. The simulation results demonstrate that the algorithm is efficient in correcting the INS estimated trajectory. In addition, we proposed an enhanced 2-layer addition to the algorithm to improve the performance when the map resolution is low. We showed by simulations that our proposed algorithm outperforms existing map matching algorithms in terms of navigational accuracy and is also robust to varying levels of sensor noise. In terms of future work, we have validated our algorithm using gravitational measurements however it can be potentially generalized to other aiding resources, such as terrain elevation and magnetic field, that have similar maps. Additionally, we currently use a straightforward method when combining the output of the algorithm with the INS position – the INS position is simply reset with the corrected position. Accordingly, another area of future work is to develop a more sophisticated fusion process.

6 Conflict of interest

None

References

Affleck, C. A. & Jircitano, A. (1990), Passive gravity gradiometer navigation system, in ‘IEEE Symposium on Position Location and Navigation’, pp. 60–66.

Bergman, N. (1997), ‘A bayesian approach to terrain-aided navigation’, IFAC Proceedings Volumes 30(11), 1457–1462.

Gade, K. (2010), ‘A non-singular horizontal position representation’, The journal of navigation 63(3), 395–417.

Groves, P. D. (2013), Principles of GNSS, inertial, and multisensor integrated navigation systems, Artech House.

Han, Y., Wang, B., Deng, Z. & Fu, M. (2016), ‘An improved TERCOM-based algorithm for gravity-aided navigation’, IEEE Sensors Journal 16(8), 2537–2544.

Han, Y., Wang, B., Deng, Z. & Fu, M. (2017), ‘A combined matching algorithm for underwater gravity-aided navigation’, IEEE/ASME Transactions on Mechatronics 23(1), 233–241.

Hemann, G., Singh, S. & Kaess, M. (2016), Long-range GPS-denied aerial inertial navigation with LIDAR localization, in ’2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)’, IEEE, pp. 1659–1666.

Hirt, C., Claessens, S., Fecher, T., Kuhn, M., Pail, R. & Rexer, M. (2013), ‘New ultrahigh-resolution picture of Earth’s gravity field’, Geophysical research letters 40(16), 4279–4283.

Hostetler, L. & Andreas, R. (1983), ‘Nonlinear Kalman filtering techniques for terrain-aided navigation’, IEEE Transactions on Automatic Control 28(3), 315–323.

Jircitano, A. & Dosch, D. (1991), Gravity aided inertial navigation system (GAINS), in ‘Proc. 47th Annual Meeting of The Institute of Navigation’, pp. 221–229.

Kamgar-Parsi, B. & Kamgar-Parsi, B. (1999), Vehicle localization on gravity maps, in ‘Unmanned Ground Vehicle Technology’, Vol. 3693, International Society for Optics and Photonics, pp. 182–191.

Lee, J., Kwon, J. & Yu, M. (2015), ‘Performance evaluation and requirements assessment for gravity gradient referenced navigation’, Sensors 15(7), 16833–16847.

Liu, F., Li, F. & Jing, X. (2019), ‘Navigability analysis of local gravity map with projection pursuit-based selection method by using gravitation field algorithm’, IEEE Access 7, 75873–75889.

Liu, F., Li, F., Lin, N. & Jing, X. (2019), ‘Gravity aided positioning based on real-time ICCP with optimized matching sequence length’, IEEE Access 7, 97440–97456.

Muniraj, D. & Farhood, M. (2017), A framework for detection of sensor attacks on small unmanned aircraft systems, in ‘2017 International Conference on Unmanned Aircraft Systems (ICUAS)’, IEEE, pp. 1189–1198.
Ren, M. & Karimi, H. A. (2009), ‘A hidden Markov model-based map-matching algorithm for wheelchair navigation’, *The Journal of Navigation* 62(3), 383–395.

Scales, K. L., Hazen, E. L., Jacox, M. G., Edwards, C. A., Boustany, A. M., Oliver, M. J. & Bograd, S. J. (2017), ‘Scale of inference: on the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data’, *Ecography* 40(1), 210–220.

Stroock, D. W. (2013), *An introduction to Markov processes*, Vol. 230, Springer Science & Business Media.

Titterton, D. & Weston, J. (2004), *Strapdown Inertial Navigation Technology*, Institution of Engineering and Technology.

Viterbi, A. (1967), ‘Error bounds for convolutional codes and an asymptotically optimum decoding algorithm’, *IEEE Transactions on Information Theory* 13(2), 260–269.

Wang, B., Yu, L., Deng, Z. & Fu, M. (2016), ‘A particle filter-based matching algorithm with gravity sample vector for underwater gravity aided navigation’, *IEEE/ASME Transactions on Mechatronics* 21(3), 1399–1408.

Wang, H., Wu, L., Chai, H., Bao, L. & Wang, Y. (2017a), ‘Location accuracy of INS/gravity-integrated navigation system on the basis of ocean experiment and simulation’, *Sensors* 17(12), 2961.

Wang, X., Gilliam, C., Kealy, A., Close, J. & Moran, B. (2022), ‘Probabilistic map matching for robust inertial navigation aiding’, *arXiv*.

Wei, E., Dong, C., Liu, J., Yang, Y., Tang, S., Gong, G. & Deng, Z. (2017), ‘A robust solution of integrated SITAN with TERCOM algorithm: weight-reducing iteration technique for underwater vehicles’ gravity-aided inertial navigation system’, *NAVIGATION, Journal of the Institute of Navigation* 64(1), 111–122.