Abstract

This paper presents a tightly integrated navigation system combining global positioning system (GPS) and inertial-based simultaneous localization and mapping (SLAM) for UAV platforms. GPS raw measurements, called pseudorange and pseudorange rate, are directly fused to an inertial SLAM filter. A compressed form of unscented filtering is implemented by partitioning the map into a local and global one. The performance of the proposed method is analysed using a high-fidelity 6-degrees-of-freedom simulator, demonstrating accurate and robust navigation even under a single satellite observation. The information gain of bearing and elevation angles is further analysed offering effective sensing strategies.

1 INTRODUCTION

With the advances in low-cost inertial sensor technology and global positioning system (GPS), the six-degrees-of-freedom (6-DOF) vehicle state can be effectively estimated by fusing GPS and inertial information, which has been a crucial step towards autonomous guidance and flight control [12, 10].

In many robotic applications however, the vehicles need to perform tasks close to environments, for example underground, urban canyon or forest, where GPS signal can be degraded or completely blocked. SLAM has been actively investigated to overcome this issue by mapping unknown environments, whilst simultaneously localising the vehicle utilising the on-line constructed map [3]. Contrast to conventional navigation systems, such as terrain and image matching systems, no infrastructure or a priori information about the environment is required, making it quite attractive for standalone or aided navigation in GPS-denied environments [11, 9, 4].

The feasibility of integrating GPS measurement and SLAM has been previously studied. [7] demonstrates a loosely-coupled integration of GPS and Inertial SLAM and [2] showed a tightly-coupled integration for a 3-DOF land vehicle model (thus non-inertial based). The work in [1] combined a compressed SLAM [5] and unscented filtering to relieve the nonlinearity in the sensing and vehicle model. In this work, we further elaborate our previous work using a full 6-DOF inertial navigation model aiming for UAV applications.

The key contributions of this paper are:

- Tightly-coupled integration of GPS and inertial SLAM using a full 6-DOF model
- Performance analysis with a varying number of satellite measurements
- Analysis of the geometrical effects of angular observation on the position solution.

The overall architecture of the fusion system is shown in Fig. 1. The measurement consists of pseudorange and pseudorange rates from visible satellite vehicles, and bearing and elevation angles from a camera. The initial positions of landmarks can be estimated by a triangulation process but in this work a range is assumed available to simplify the implementation of the method. The map database is maintained within the fusion filter by storing estimated landmark positions and orbital ephemeris of the satellite vehicles.

The remainder of the paper is outlined as follows. In Section 2, we review the unscented filtering method. We then provide the integration of GPS and IMU-based SLAM in an unscented filtering framework in Section 3, and extend it to compressed filter form in Section 4. Experimental results will be shown in Section 5 with further analysis on the geometrical effects of angular measurement in Section 6, followed by Conclusion.

2 Background on Unscented filtering

The unscented filter has been proposed for the nonlinear state estimation based on the unscented transformation, in which multivariate moment integrals can be numerically approximated [6, 1]. In contrast to the EKF, the
corresponding weights $w_i$ to scale the sigma points, commonly a design parameter to capture the distribution of the state vector. Then these points are used for nonlinear transformation.

At the prediction step, first an augmented vector is formed comprising the state vector $\hat{X}_{k|k}$ and the control input with an augmented covariance matrix $\hat{P}_{k|k}$

$$\hat{X}_{k|k} = \begin{bmatrix} \hat{X}_{k|k} \\ \hat{U}_{k} \end{bmatrix}, \quad \hat{P}_{k|k} = \begin{bmatrix} P_{k|k} & 0 \\ 0 & Q_k \end{bmatrix}$$ (1)

Then the sigma points $\{\chi_{i|k}^j\}_{i=0}^{2n}$ are selected from the augmented vector, with $n$ being $\dim(x) + \dim(u)$, and corresponding weights $w_i$,

$$\begin{align*}
\chi_{0|k}^j &= \hat{X}_{k|k}, \quad w_0 = \frac{n}{n + \kappa} \\
\chi_{i+n|k}^j &= \hat{X}_{k|k} + \sqrt{(n + \kappa)P_{k|k}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad w_i = \frac{1}{2(n + \kappa)} \\
\chi_{i|k}^j &= \hat{X}_{k|k} - \sqrt{(n + \kappa)P_{k|k}} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad w_{i+n} = \frac{1}{2(n + \kappa)}
\end{align*}$$

(2)

where $[A]_i$ represents the $i^{th}$ column of a matrix $A$. $\kappa$ is a design parameter to scale the sigma points, commonly selected so that the scale parameter $n + \kappa = 3$ for Gaussian distributions [8, 6]. The mean $\hat{X}_{k+1|k}$ and covariance matrix $\hat{P}_{k+1|k}$ of the predicted state are computed using the predicted sigma points $\chi_{i+1|k}^j = f(\chi_{i|k}^j)$,

$$\hat{X}_{k+1|k} = \sum_{i=0}^{2n} w_i \chi_{i+1|k}^j$$ (3)

$$\hat{P}_{k+1|k} = \sum_{i=0}^{2n} w_i \left( \chi_{i+1|k}^j - \chi_{0|k}^j \right) \left( \chi_{i+1|k}^j - \chi_{0|k}^j \right)^T$$

At the measurement update stage, a set of sigma points $\{\chi_{i+k+1|k}^j\}_{i=0}^{2n}$ is selected from the predicted state (typically the control input is not related to measurement and thus not included in sampling) and is used to predict the sigma points $Z_{k+1}^i = h(\chi_{k+1|k}^j)$. The predicted measurement and related cross-correlation and covariance matrices are computed accordingly

$$\hat{Z}_{k+1} = \sum_{i=0}^{2n} w_i Z_{k+1}^i$$ (4)

$$\begin{align*}
\hat{P}_{zz,k+1} &= \sum_{i=0}^{2n} w_i (Z_{k+1}^i - \hat{Z}_{k+1}) (Z_{k+1}^i - \hat{Z}_{k+1})^T \\
\hat{P}_{xz,k+1} &= \sum_{i=0}^{2n} w_i (\chi_{k+1|k}^j - \chi_{0|k}^j) (Z_{k+1}^i - \hat{Z}_{k+1})^T
\end{align*}$$

Finally, the state and covariance matrix can be updated as follows

$$\begin{align*}
S_{k+1} &= P_{zz,k+1} + R_{k+1} \\
K_{k+1} &= P_{xz,k+1} S_{k+1}^{-1} \\
\hat{X}_{k+1|k+1} &= \hat{X}_{k+1|k} + K_{k+1} (Z_{k+1} - \hat{Z}_{k+1}) \\
P_{k+1|k+1} &= P_{k+1|k} - K_{k+1} S_{k+1} K_{k+1}^T
\end{align*}$$

If a new landmark $l_k$ is observed, it should be initialized and augmented through a non-linear initialization function $g(\cdot)$ which is the inverse process of the measurement function $h(\cdot)$ with a noise $\nu_k$,

$$l_k = g(x_k, z_k, \nu_k).$$ (6)

### 3 Unscented Inertial SLAM

The full state vector is partitioned into a local map and a global map with the estimated state vector and covariance matrix being

$$\begin{bmatrix} \hat{x}_L \\ \hat{m}_L \end{bmatrix} = \begin{bmatrix} \hat{x}_G \\ m_G \end{bmatrix}$$

(7)

$$\begin{bmatrix} P_L \\ P_{LG} \\ P_{GL} \\ P_{GG} \end{bmatrix} = \begin{bmatrix} P_{xx} \\ P_{xm} \\ P_{xG} \\ P_{mG} \end{bmatrix}$$

(8)

where the local state $x_L$ includes the vehicle state $x$ and the local landmarks $m_L$, while other landmarks are stacked in $m_G$.

The unscented filter in SLAM utilises a partial sampling scheme as the nonlinearity appears only in the vehicle state transition, whilst the map state having linear dynamics. The equivalence of partial sampling and full sampling scheme was provided in [1] showing that the performance of the partial sampled filter is not affected by the number of feature landmarks.

The state vector $x_L$ consists of the vehicle state (INS state and receiver clock state), the map state with the control input from IMU measurement,

$$x_k = \begin{bmatrix} p^T, v^T, \phi^T, b_0^T, b_γ^T, c_t, c_d \end{bmatrix}^T$$ (9)
\[ \mathbf{m}_k = [\mathbf{m}_1^T, \mathbf{m}_2^T, \ldots, \mathbf{m}_L^T]^T \tag{10} \]
\[ \mathbf{u}_k = [\mathbf{f}^T, \mathbf{\omega}^T]^T, \tag{11} \]

with

- Vehicle position \( \mathbf{p} = [x, y, z]^T \)
- Vehicle velocity \( \mathbf{v} = [v_x, v_y, v_z]^T \)
- Euler angles \( \mathbf{\phi} = [\phi, \theta, \psi]^T \)
- Accelerometer bias \( \mathbf{b}_a = [b_{a_x}, b_{a_y}, b_{a_z}]^T \)
- Gyroscope bias \( \mathbf{b}_g = [b_{g_x}, b_{g_y}, b_{g_z}]^T \)
- Receiver clock bias \( c_t b \) with \( c \) being the speed of light
- Receiver clock drift \( c_t d \)
- Map position \( \mathbf{m}_i = [m_{ix}, m_{iy}, m_{iz}]^T \)
- IMU control input \( \mathbf{u}_k = [\mathbf{f}, \mathbf{\omega}] \).

### 3.1 Prediction

The state transformation function can be decoupled into a nonlinear part \( \hat{\mathbf{x}}_{k+1}|k = \mathbf{f}(\hat{\mathbf{x}}_k|k) \) and a linear part \( \hat{\mathbf{m}}_{k+1} = \hat{\mathbf{m}}_k \), thus giving

\[
\begin{bmatrix}
\hat{p}_{k+1} \\
\hat{\mathbf{v}}_{k+1} \\
\hat{\mathbf{b}}_{a,k+1} \\
\hat{\mathbf{b}}_{g,k+1} \\
\hat{c}_t b_{k+1} \\
\hat{c}_t d_{k+1} \\
\hat{\mathbf{m}}_{k+1}
\end{bmatrix} =
\begin{bmatrix}
\hat{p}_k + \hat{\mathbf{v}}_k \Delta T \\
\hat{\mathbf{v}}_k + C_{\mathbf{v}, k}(\hat{\mathbf{f}}_k + \hat{\mathbf{b}}_{a,k}) \Delta T + g^n \Delta T \\
\hat{\mathbf{b}}_{a,k} + \hat{\mathbf{b}}_{g,k} \\
\hat{\mathbf{b}}_{a,k} + \hat{\mathbf{b}}_{g,k} \\
\hat{\mathbf{c}}_t b_k + \hat{\mathbf{c}}_t d_k \Delta T \\
\hat{\mathbf{c}}_t d_k \\
\hat{\mathbf{m}}_k
\end{bmatrix},
\tag{12}
\]

where \( \Delta T \) is the prediction sampling time and \( g^n \) the gravitational vector \( [0; 0; 9.8] \) \((m/s^2)\).

For the control augmented vector, the sigma points \( \{\chi_k|^i\}_{i=0}^{2n_x} \) are selected with corresponding weights \( \omega_i \), where \( n = \text{dim}(x) + \text{dim}(u) = 23 \). The distribution of the predicted state \( \chi_{k+1}|k \) can then be calculated according to (3). The cross-correlation \( P_{x_m,k+1|k} \) between \( \chi_{k+1}|k \) and \( \mathbf{m}_{k+1}|k \) needs to be computed as in \( \text{[1]} \),

\[
P_{x_m,k+1|k} = \left( \sum_{i=0}^{2n_x} \omega_i \chi_k|^i \hat{\chi}_k|^i, T \mathbf{P}^{-1} \hat{\chi}_k|^i, k\right) \mathbf{P}_{x_m,k|k}
\]

\[
\triangleq \mathbf{F}_{k+1} \mathbf{P}_{x_m,k|k} \tag{13}
\]

where \( \chi_k|^i = \chi_{k+1}|k - \chi_k|^i \) and \( \mathbf{F}_{k+1} \) is equivalent to the inferred Jacobian as in \( \text{[6]} \).

### 3.2 Update

The measurement consists of either GPS pseudorange and pseudorange rate measurement from a satellite vehicle or a range/bearing/elevation measurement from a landmark

\[
z_k = \begin{cases} 
\hat{z}_k^i = \begin{bmatrix} \hat{\rho}_k^1 & \hat{\rho}_k^2 \end{bmatrix}, & \text{for } i^{th} \text{- Satellite Vehicle, or} \\
\hat{z}_k^j = \begin{bmatrix} \hat{\rho}_k^j & \hat{\phi}_k^j & \theta_k^j \end{bmatrix}, & \text{for } j^{th} \text{- Landmark}
\end{cases}
\]

where pseudorange \( \hat{\rho}_k^i = \| \mathbf{P}_{SV} \mathbf{p}_k - \mathbf{p}_k \|_2 + c t_b_k + v_p \), pseudo-range rate \( \hat{\rho}_k^j = \| \mathbf{V}_{SV} \mathbf{v}_k - \mathbf{v}_k \|_2 + c t_d_k + v_p \) and \( (\hat{\rho}_k^j, \hat{\phi}_k^j, \hat{\theta}_k^j) \) is a range, bearing, and elevation measurement.

Using the drawn sigma points, the predicted measurement and its covariance can be computed as in Equation 4. The cross-correlation between the state and measurement can be directly computed using the sigma points

\[
P_{x_z,k+1} = P_{x_z,k|k} \left( \sum_{i=0}^{2n_x} \omega_i \hat{z}_i|^i, k \right) \mathbf{P}_{x_z,k|k}
\]

\[
\triangleq \mathbf{P}_{x_z,k+1} \mathbf{H}_{k+1}^T
\tag{14}
\]

where \( \hat{z}_i|^i, k = \hat{z}_i|^i, k - \hat{z}_i^0|^i, k \) and \( \mathbf{H}_{k+1} \) is an inferred measurement Jacobian from the sigma samples.

### 3.3 Initialisation

The sigma points \( \chi_k|^i \) are used to predict the landmark position \( \{\chi_k|^i\}_k \) and the cross-correlations with the vehicle and map are computed respectively as

\[
P_{x_z,l,k} = \left( \sum_{i=0}^{2n_x} \omega_i \hat{z}_i|^i, k \right) \mathbf{P}_{x_z,l,k|k}
\]

\[
\triangleq \mathbf{G}_k \mathbf{P}_{x_z,l,k|k} \tag{15}
\]

\[
P_{m,l,k} = \mathbf{G}_k \mathbf{P}_{x_m,l,k|k} \tag{16}
\]

where \( \mathbf{G}_k = \mathbf{L}_i|^i, k - \mathbf{L}_i^0|^i, k \).

### 4 Compressed Implementation

In the local area, we can efficiently predict \( \hat{\mathbf{x}}_{L,k+1|k} \) and \( \mathbf{P}_{L,k+1|k} \) in real-time using the local inferred Jacobian (termed a local linear factor) \( \mathbf{F}_{k+1} \) for the propagation of \( \mathbf{P}_{x_m,k+1|k} \) as in (13). Correspondingly, the global linear factor \( \mathbf{F}_{G,k+1} \) for the prediction of \( \mathbf{P}_{G,k+1} \) can be extended as

\[
\mathbf{F}_{G,k+1} = \begin{bmatrix} \mathbf{F}_{k+1} & 0 \\ 0 & \mathbf{I}_{n_G} \end{bmatrix} \tag{17}
\]

For the \( j^{th} \) landmark observation, we can get a local linear factor \( \mathbf{H}_{k+1} = [\mathbf{H}_{x,k+1}, \mathbf{H}_{m,k+1}] \) as to (14), where \( \mathbf{H}_{x,k+1} \) and \( \mathbf{H}_{m,k+1} \) correspond to the vehicle state and
the $j^{th}$-landmark, respectively. The global linear factor $H_{G,k+1}$ for the whole map can then be computed as

$$H_{G,k+1} = [H_{x,k+1}, 0, \cdots, 0, H_{m,j,k+1}, 0, \cdots, 0].$$

(18)

Similarly, for the landmark initialization step, we can get the local linear factor $G_{k+1}$ through (15), and the global linear factor $G_{G,k+1}$ becomes

$$G_{G,k+1} = \begin{bmatrix} G_{k+1} & 0_{mG} \end{bmatrix}$$

(19)

As described in [5], when the vehicle runs in the local area, the local prediction and update information of $P_{LG}$, $P_{GG}$ and $m_{G}$ are preserved in three auxiliary variables, namely $\alpha$, $\beta$ and $\gamma$. By using the foregoing linear factors, the full map update can be performed.

5 Experimental Results

Using a high-fidelity simulator GPSoft, GPS raw measurements are generated with the pseudorange noise having a standard deviation of $2m$ and pseudorange rate error with $1m/s$. The output rate of both measurements is $1Hz$. The IMU output is $100Hz$ with accelerometer bias of $2.0 \times 10^{-2}m/s^2$ and gyroscope bias of $100^\circ/hr$ (or $2.8^\circ/sec$). This noise levels correspond to typical low-quality and low-cost inertial sensors. For the landmark observation, 400 landmarks are simulated with a $10Hz$ output rate and a $\pm20^\circ$ down-looking field of view.

To verify the benefits of tightly-coupled integration, the number of satellite vehicles reduces from seven to three (at $50sec$) to one (at $100sec$) which is shown as a dashed line in Fig. 2.

Figure 2(top) shows the estimated 3D trajectory and Fig. 2(bottom) depicts the evolution of the map uncertainties plotted with the number of satellite vehicles on the same $y$-axis. It can be observed that when the number of SVs dropped from seven to three, the newly initialized landmarks show larger errors due to reduced information from the SVs. The second group of large landmark uncertainties from around $140sec$ are related to the second circuit of the trajectory. It can be seen however, that those uncertainties from the two circuits approach to a similar lower bound which is determined by the sensing uncertainties. Figure 4 depicts the estimated biases from the accelerometers and gyroscopes, showing that the estimated biases converge reliably to the true biases. The convergence of the receiver clock bias and drift is also shown in Fig. 4. Due to the constant drift term, the clock bias changes linearly. It can be seen that SLAM observation can reliably and effectively estimate the receiver clock errors when the number of satellite vehicle is only one after $100sec$, which is crucial for precise ranging and timing outputs from the integrated system.

6 Geometrical Effects of Bearing/Elevation Measurements

Understanding the effects of the sensing geometry of satellite vehicles and ground landmarks can help select a most informative set of satellite/landmarks. One of key questions can be: given a generated SLAM map, what sensing geometry of bearing and elevation observations can maximise the information gain on the position solution?

For the purpose of analysis, we consider a simple case as illustrated in Fig.5, in which a vehicle flies a straight-line trajectory with a constant height, observing two landmarks on the ground. The information gain from the sensing is $I \triangleq H^TH$ using a linearised observation matrix $H$. If $H$ is an invertible $3 \times 3$ matrix, then $\det(H^TH) = (\det(H))^2$. Thus we investigate $\det(H)$ to gain intuitive understanding. Given the vehicle position $x = [x, y, z]^T$ and $i^{th}$-landmark position $[x_i, y_i, z_i]^T$, ...
the bearing $\phi_i$ and elevation $\theta_i$ observations are

$$\phi_i = \arctan \left( \frac{y_i - y}{x_i - x} \right)$$

(20)

$$\theta_i = \arctan \left( \frac{z_i - z}{\sqrt{(x_i - x)^2 + (y_i - y)^2}} \right).$$

(21)

It can be seen that 3 observations are required to solve for 3 unknowns: $x$, $y$ and $z$ position. Thus we stack 3 angle observations $z = [\phi_1, \phi_2, \theta_1]^T$ and by linearisation and trigonometric manipulation, the absolute value of the determinant $|\det(H)|$ can be shown as

$$|\det(H)| = \left| \frac{\sin(\Delta \phi_{12})}{r_{1o}r_{2o}} \right| \left| \frac{\cos(\theta_1)}{r_1} \right|,$$

(22)

where $\Delta \phi_{12}$ is the relative bearing angle between two landmarks, with $r_{1o}, r_{2o}$ being the horizontal distance to each landmark, and $r_1$ the full range to the landmark 1. It can be observed that

- The relative bearing angle contributes to the horizontal information gain, while the elevation to the vertical information.
- When $\Delta \phi_{12} = \pm 90^\circ$ the sine term becomes maximum and as the vehicle moves closer to the landmarks, the horizontal gain increases. If the vehicle position is along the line connecting two landmarks, the horizontal position becomes unobservable which corresponds to the condition of $\Delta \phi_{12} = 0^\circ$ or $180^\circ$.
- The vertical gain proportional to the cosine of the elevation and inverse range to the landmark. The maximum gain can be achieved with $\pm 45^\circ$ elevation condition. This is illustrated in Figure 6, in which the height is set to 50m and a landmark is located at [50, 0]m. Note the vertical gain drops to zero as the vehicle flies directly over the landmark, which corresponds to the 90$^\circ$ elevation angle and the vertical position becomes unobservable. As the vehicle moves away from the landmark, the vertical gain approaches to zero, which is related to a lower elevation angle.
7 Conclusions

In this paper, we presented a tightly-coupled integration of GPS and inertial-based SLAM using a full 6-DOP inertial navigation model, extending the previous work on a 3-DOP model. The simulation results demonstrated the benefits of fusing GPS pseudorange and pseudorange rates with inertial SLAM, enabling a reliable and robust navigation solution even with a single satellite vehicle, and the fast convergence of IMU and receiver biases. In addition, the effects of sensing geometry of bearing and elevation were analysed, showing that ±90° relative bearing of two landmarks can maximise the horizontal information gain, whilst ±45° elevation angle can maximise the vertical gain. Currently real flight data is being examined to verify the performance in a real environment.

References