Compressed Pseudo-SLAM: Pseudorange Integrated Generalised Compressed SLAM

Jonghyuk Kim, Hongkyoon Byun
University of Technology Sydney[1]

Jose Guivant
University of New South Wales[2]

Tor Arne Johansen
Norwegian University of Science and Technology, Norway[3]

Abstract

This paper addresses the fusion of the pseudorange/pseudorange rate observations from global navigation satellite system (GNSS), and the inertial-visual simultaneous localisation and mapping (SLAM) to achieve reliable navigation of unmanned aerial vehicles (UAVs). This work extends the previous work on a simulation-based study [Kim et al.(2017)], and evaluates the method to a flight dataset collected from a fixed-wing UAV platform. We propose to use the generalised compressed filter which can effectively accumulate the information gain acquired from a local map, and update the global map in a much lower rate. The fusion filter also models and estimates the receiver clock and drift, which is crucial to integrate the pseudorange and pseudorange rate measurements. Evaluation results will show that the horizontal navigation error is effectively constrained even with 1 satellite vehicle and 1 landmark observations, thanks to the direct fusion of pseudorange and vision data.

1 Introduction

Autonomous unmanned aerial vehicles (UAVs) have attracted much attention from both industries and defence over the last several decades. With the advances in the low-cost inertial sensor technology and global navigation satellite system (GNSS), the navigation solution (or position, velocity, attitude and time) can be accurately estimated by integrating the information. This data fusion approach has enabled the higher-level autonomy, such as intelligent path-planning and decision-making [Parkinson et al.(1997)], [Skulstad et al.(2015)].

Simultaneous localisation and mapping (SLAM) has been quite successful in delivering navigational solutions. For example, [Kim and Sukkarieh(2007)] demonstrate the integration of inertial measurement unit (IMU) and inertial navigation system (INS) as the core odometry together with a visual perception sensor on a UAV platform. [Nützi et al.(2011)] integrated monocular vision and inertial sensor on a drone platform addressing the monocular scale-ambiguity problem. [Li and Mourikis(2013)] showed visual-inertial odometry and SLAM by utilising a sliding-window based bundle-adjustment optimisation on a 6-degrees-of-freedom vehicle. [Sjanic et al.(2017)] applied the expectation-maximisation technique for visual-inertial navigation of a fixed-wing flight data. Recently, [Vidal et al.(2018)] demonstrated a high-speed drone by combining an event camera and inertial observations.

Although the abovementioned approaches have been quite successful particularly for low-dynamic platforms such as drones, SLAM for high-speed UAV platforms has had limited success, largely due to the low-quality inertial odometry and the high computational complexity inherent to SLAM. If the number of landmark observations is low (in this work it is less than 3 landmarks per camera frame on average, as illustrated in Fig. 1), the inertial odometry can drift between the re-observations of the SLAM map.

To improve the inertial SLAM, we integrate the raw pseudorange/pseudorange rates and the inertial SLAM system. Although integrating the global navigation information can be counter motivational to SLAM, it is mostly or partially available to aerial vehicles operating near-Earth environments, and can aid the inertial odometry effectively. To further improve the computational efficiency, we adopt the generalised compressed filtering method [Guivant(2017)], in which the local information gain is accumulated and propagated a much lower rate. This work extends the previous work on a simulation-based study by the authors. The contributions of the work are:

- Compressed Pseudo-SLAM: pseudorange integrated SLAM using a generalised compressed filtering. As

*[1] {jonghyuk.kim, hongkyoon.byun}@uts.edu.au, [2] j.guivant@unsw.edu.au, and [3] tor.arne.johansen@ntnu.no

depicted in Figure 1(b), the number of visual observations is sparse during this flight and only available when the vehicle reaches at certain altitude with clear view. This necessitates the fusion of pseudorange to aid the on-board IMU.

- Evaluation of the method using a real flight dataset.

The remainder of the article is outlined as follows. Section 2 provides the system models for the nonlinear state transition and observation used in this work, and Section 3 details the generalised compressed SLAM with a compressed update of the local and global map. Experimental results are provided in Section 4, analysing the filter performance and the computational complexity. Section 5 will conclude with future direction.

2 System Models

A continuous-time stochastic dynamic system with a nonlinear transition model $\mathbf{f}(\cdot)$ and an observation model $\mathbf{h}(\cdot)$ can be written as

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), \mathbf{w}(t)) \quad (1)$$

$$\mathbf{z}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{v}(t)), \quad (2)$$

where $\mathbf{x}(t)$ and $\mathbf{u}(t)$ are the state and control input vectors at time t, respectively, with $\mathbf{w}(t)$ being the process noise with a noise strength matrix $\mathbf{Q}$, and $\mathbf{z}(t)$ is the observation vector with $\mathbf{v}(t)$ being the observation noise with a strength matrix of $\mathbf{R}$.

In the inertial-based SLAM, the state vector consists of a vehicle state which represents the kinematics of the inertial navigation, inertial sensor errors, and map landmark states. In this work, a 17-state model is designed to estimate the vehicle position and attitude (6), internal velocity and sensor biases (9), receiver clock states (2), as well as the variable-size map states. The control input is the IMU measurements and drives the inertial kinematic equations. Thus the state vector $\mathbf{x} = (\mathbf{x}_v^T, \mathbf{m}^T)^T$ becomes,

$$\mathbf{x}_v = [\mathbf{p}^{nT}, \mathbf{v}^{nT}, \boldsymbol{\psi}^{nT}, \mathbf{b}_a^{bT}, \mathbf{b}_g^{bT}, ct_b, ct_d]^T \quad (3)$$

$$\mathbf{m} = [\mathbf{m}_1^{nT}, \mathbf{m}_2^{nT}, \cdots, \mathbf{m}_N^{nT}]^T \quad (4)$$

$$\mathbf{u} = [\mathbf{f}^{bT}, \boldsymbol{\omega}^{bT}]^T, \quad (5)$$

where

- Vehicle position in the navigation frame[1] $\mathbf{p}^n = [x, y, z]^T$
- Vehicle velocity $\mathbf{v}^n = [v_x, v_y, v_z]^T$
- Vehicle attitude (Euler angles in roll, pitch and yaw) $\boldsymbol{\psi}^n = [\phi, \theta, \psi]^T$
- Accelerometer bias in the body frame[2] $\mathbf{b}_a^b = [b_{ax}, b_{ay}, b_{az}]^T$
- Gyroscope bias $\mathbf{b}_g^b = [b_{gx}, b_{gy}, b_{gz}]^T$
- Receiver clock bias ct_b with c being the speed of light
- Receiver clock drift ct_d
- Map position $\mathbf{m}_i^n = [m_{ix}, m_{iy}, m_{iz}]^T$
- Accelerometer measurement $\mathbf{f}^b = [f_x, f_y, f_z]^T$
- Gyroscope measurement $\boldsymbol{\omega}^b = [\omega_x, \omega_y, \omega_z]^T$.

2.1 Dynamic Model

The system dynamic model consists of the kinematic equations of the inertial navigation driven by the IMU measurements which are the specific force (or the sum of dynamic acceleration of the vehicle and gravity) and angular rate. The IMU bias errors are modelled as a random walk process, while the map states are modelled as a random constant due to their stationary nature,

$$\begin{aligned}
&\dot{\mathbf{p}}^n = \mathbf{v}^n \\
&\dot{\mathbf{v}}^n = \mathbf{C}_b^n(\mathbf{f}^b - \mathbf{b}_a^b) - 2\boldsymbol{\omega}_{ie}^n \times \mathbf{v}^n + \mathbf{g}^n(\mathbf{p}^n) \\
&\dot{\boldsymbol{\psi}}^n = \mathbf{E}_b^n(\boldsymbol{\omega}^b - \mathbf{b}_g^b) \\
&\dot{\mathbf{b}}_a^b = \mathbf{w}_a, \quad \dot{\mathbf{b}}_g^b = \mathbf{w}_g \\
&\dot{ct}_b = ct_d, \quad \dot{ct}_d = w_d \\
&\dot{\mathbf{m}}_{1\cdots N}^n = \mathbf{0},
\end{aligned}$$

where

- $\boldsymbol{\omega}_{ie}^n$ is the Earth rotation rate in the navigation frame
- $\mathbf{g}^n(\mathbf{p}^n)$ is the gravitational acceleration
- $\mathbf{w}_a, \mathbf{w}_g, w_d$ are noise processes for accelerometers, gyroscopes, and receiver clock drift respectively
- $\mathbf{C}_b^n$ and $\mathbf{E}_b^n$ are the direction cosine matrix transforming a vector from the body- to the navigation-frame, and the matrix transforming a body rate to an Euler angle rate, respectively,

$$\mathbf{C}_b^n = \begin{bmatrix} c_\theta c_\psi & -c_\phi s_\psi + s_\phi s_\theta c_\psi & s_\phi s_\psi + c_\phi s_\theta c_\psi \\ c_\theta s_\psi & c_\phi c_\psi + s_\phi s_\theta s_\psi & -s_\phi c_\psi + c_\phi s_\theta s_\psi \\ -s_\theta & s_\phi c_\theta & c_\phi c_\theta \end{bmatrix}$$

$$\mathbf{E}_b^n = \begin{bmatrix} 1 & s_\phi t_\theta & c_\phi t_\theta \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi / c_\theta & c_\phi / c_\theta \end{bmatrix},$$

where $s_{(\cdot)}$, $c_{(\cdot)}$, and $t_{(\cdot)}$ are shorthand notations for $\sin(\cdot)$, $\cos(\cdot)$, and $\tan(\cdot)$, respectively.

[1] A superscript-n denotes a navigation frame. In this work, MGA94 (Map Grid of Australia 1994) is used as a local-fixed, local-tangent frame with North, East and Down (NED) (m)

[2] A superscript-b denotes a body frame attached to the vehicle.

2.2 Observation Model

In all-source aiding strategy, the observations consist of all-available aiding information which are, in this work, a set of pseudoranges (ρ) and pseudorange rates ($\dot{\rho}$ measured from a GNSS receiver, and/or a set of bearing ϕ, elevation θ, and/or range r observations from a camera sensor in which the range is estimated from the known visual landmark size (white plastic sheets of $1m \times 1m$),

$$\rho_i = \sqrt{(X_i - x)^2 + (Y_i - y)^2 + (Z_i - z)^2} + ct_b + v_\rho$$
$$\dot{\rho}_i = l_{xi}(V_{xi} - v_x) + l_{yi}(V_{yi} - v_y) + l_{zi}(V_{zi} - v_z) + c\dot{t}_b + v_{\dot{\rho}}$$
$$r_j = \sqrt{(m_{jx} - x)^2 + (m_{jy} - y)^2 + (m_{jz} - z)^2} + v_r \quad (6)$$
$$\phi_j = \arctan \frac{(m_{jy} - y)}{(m_{jx} - x)} + v_\phi \quad (7)$$
$$\theta_j = \arctan \frac{(m_{jz} - z)}{\sqrt{(m_{jx} - x)^2 + (m_{jy} - y)^2}} + v_\theta, \quad (8)$$

where (X_i, Y_i, Z_i) denotes the i^{th} satellite vehicle's position computed from the ephemeris data, which is converted from the ECEF (Earth-Centred, Earth-Fixed) frame to the navigation frame, with v_ρ being the noise process, and (l_{xi}, l_{yi}, l_{zi}) is the line-of-sight vector of the i^{th} SV seen from the vehicle which projects the relative velocity along the line-of-sight direction. For the range, bearing and elevation measurement, we first compute the relative position vector $((m_{jx} - x), (m_{jy} - y), (m_{jz} - z))$ of the j^{th}-landmark seen from the vehicle, and then transform it to the polar coordinate quantities. $(v_\rho, v_{\dot{\rho}}, v_r, v_\phi, v_\theta)$ represent the observation noise processes. A new landmark position is initialized by utilising the inverse function of Eqs. (6) to (8) with the estimated vehicle state and observation information.

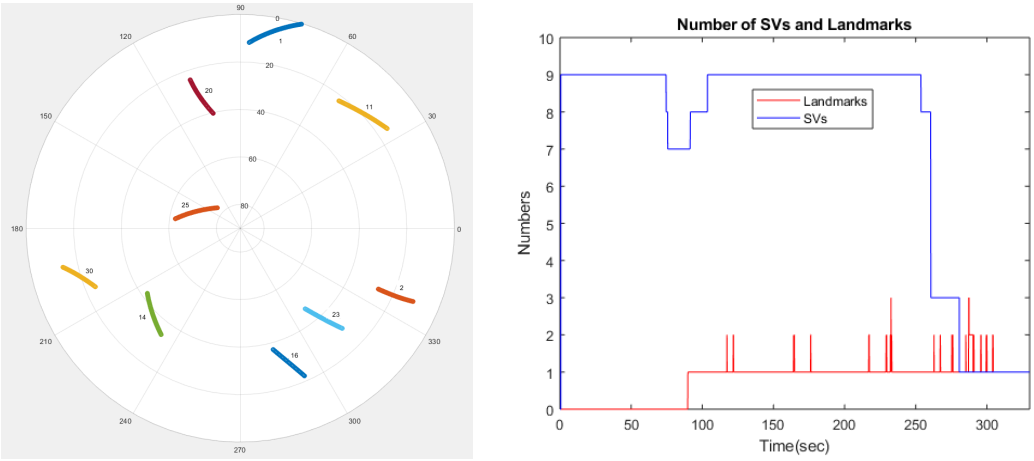
3 Generalised Compressed Filter

Using the state and observation models described in the previous section, and their corresponding linearised state and observation models, a *standard* extended Kalman filter can predict and update the state and covariance matrix as detailed in [Kim and Sukkarieh(2007)]. In a *generalised compressed* filtering framework [Guivant(2017)], the state vector is still partitioned into 1) a local map state including the vehicle state and local map, and 2) a global map for the remainder landmarks. However, a clone of the vehicle state is augmented to the state vector whenever the local map boundary changes. This augmented clone state is used to compute the information gain accumulated within the local region, which in turn is used to propagate the information to the global state. Compared to the *standard* compressed filter, in which the correlation between the local and global states was calculated explicitly using a closed-form expression, the generalised method is based on the Bayesian framework and can be used with various local filters, as well as being easily extended to multiple vehicles applications. The underlying assumption is that the evolution of correlation is *only* dependent on the local states, thus enabling the accumulation (thus compression) during the local filtering cycles. Consequently, compressed (accumulated) correlation enables the global updates at a much lower rate. This assumption is quite valid for the downward-looking camera configuration as in this work where the camera field-of-view naturally defines the boundary of the local map. During the local processing, the information gain is the information increment with respect to the initial clone state (modelled as a random constant). This information gain is then used to update the global map state, whenever the local region boundary changes.

4 Experimental Results

A flight dataset recorded from a UAV platform [Kim and Sukkarieh(2007)] is used to verify the method. An IMU sensor was from Inertial Science and is a low-grade sensor with a gyroscope bias stability of $10^\circ/hr$, delivering the specific force and angular velocity at 400 Hz. A GPS receiver from Canadian Marconi Communication (CMC, now Novatel) was used to record the pseudorange and pseudorange rate measurements at 1 Hz, as well as the satellite ephemeris data. The receiver provides integrated-carrier-phase (ICP) measurements which are used to compute the pseudorange rates. A Sony monochrome camera (with a frame rate of 25 Hz) was installed in a down-looking configuration. Artificial visual landmarks, white plastic sheets of $1m \times 1m$ to estimate the depth, were installed on the ground, and those positions were surveyed using a real-time-kinematic GPS receiver. An intensity-based fast detection algorithm is used for real-time processing on a PC104 embedded computer. Due to the known size of the sheet, the number of pixels provides the depth information. The detected landmarks are recognised within the SLAM filter using the joint-compatibility branch and bound algorithm which can effectively match a set of landmarks to the whole SLAM map. In this work, however, the number of landmarks detected in each image frame is quite low $(1-2)$, and thus the recognition performance is close to the nearest neighbour (NN) method.

The flight segment used has around 6 minutes of duration from the taking off along multiple racehorse tracks (each track is about 5 km). The average flight speed is around 120 km/hr, and the altitude is maximum 150 metres above the ground. Figure 1(a) shows the sky plot of the visible SVs in the azimuth-elevation form showing 9 visible SVs. Figure 1(b) shows the change of the number of satellite vehicles during the flight, which is simulated to drop to 3 and then 1 from 250 seconds to test the performance of the system. The number of visual landmarks



(a) (b)

Figure 1: (a) The sky view plot of the satellite vehicle (SVs) during the flight experiment. (b) A dropout of the satellites from 250 seconds was simulated together with the number visual landmarks observed.

(in red) is also shown in the figure which is $1-2$ on average. These sparse visual observations make the height aiding of the inertial navigation system quite challenging, due to the extensive range errors in the landmark observation. In this work, we stabilise the vertical height using the barometric-pressure altitude information. The absolute pressure measurements from the air data system are converted to barometric height and climb-rate and then used to aid the vertical inertial outputs using a simple α-β filter. Although this is a suboptimal approach, it simplifies the SLAM filter-tuning process with reliable performance, and thus adopted in this work.

Figures 2(a) and 2(b) show more detailed results on the estimated trajectory and the map compared to the reference trajectory and the surveyed map. The estimated trajectory from the Pseudo-SLAM is almost identical to that of the full-rate SLAM (without compression), while there is some mismatch with the on-board solution. This error seems to be the differences in processing the pseudorange measurements within the CMC-Novatel GPS receiver and the Pseudo-SLAM. A total of 85 landmarks are registered in the Pseudo-SLAM. The estimated map and uncertainty (10σ is used for presentation) are matched well to the surveyed map, although the map covariance seems over-confident with some biases. The poor visual-depth information seems to contribute to the bias errors in the estimated landmark positions, as well as the use of different GNSS receiver (Novatel RTK) for the landmark survey. Further work on SLAM filter tuning can address these errors and the over-confident uncertainty. Figures 2(c) and 2(d) present the estimated clock bias error and the gyroscope bias error along the x-axis with the estimated uncertainty, respectively. It can be observed that the clock bias error start drifting when the number of SVs drops to 3 and 1. Thanks to the subsequent visual-aiding, the clock bias error is adequately constrained within the uncertainty bound.

4.1 Computational Complexity

Figure 3(a) presents the change of the total number of landmarks (in blue) and the number of local landmarks (in red) registered in the Pseudo-SLAM filter. The dimension of the final state vector is 272, consisting of the vehicle state (17) and the landmark map (255 corresponding to 85 landmarks with 3 dimensional position). It can be seen that the number of local landmarks is less than 20 throughout the flight, although it increases towards the end of the flight. This increase was due to the higher altitude of the vehicle resulting in a bigger local map, and thus more local landmarks.

Figure 3(b) compares the computational times of the filter update from the Full-SLAM (without compression) (in red) and Pseudo-SLAM (in blue), showing the improved computational performance of the Pseudo-SLAM. The computational complexity of the Full-SLAM update is $O(N^2)$ with N being the total number of landmarks, while the Pseudo-SLAM has the complexity of $O(L^2)$ with L being the local map size, which is less than the total map size. It can be seen that the pseudorange update times are nearly constant with 1 milliseconds on average, while the local landmark update times increase towards 1.5 ms. The local-go-global updates, however, sometimes show peaks during the map transitions, which was affected by the additional data association and sorting process. During the last 60 seconds of the GNSS drop-out period, the update time is dominated by the visual updates. This result confirms that the effectiveness of the compressed filtering, and it is suitable for the

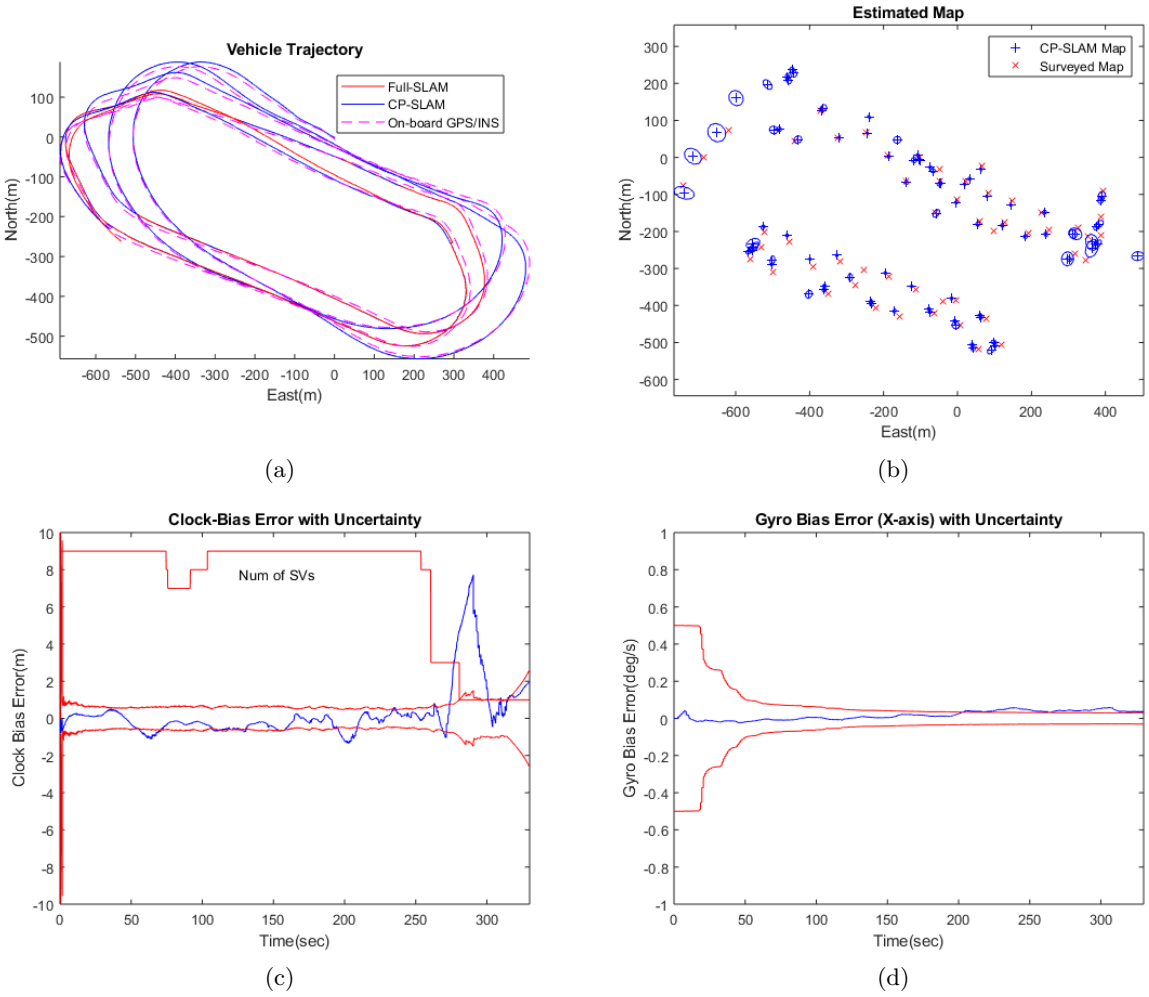


Figure 2: (a) The estimated vehicle trajectory, (b) map with uncertainty, (b) receiver clock-bias error with uncertainty, and (d) x-gyro bias error with uncertainty. The Pseudo-SLAM trajectory is compared with the full-SLAM (thus no compression) and the on-board loosely-coupled GPS/INS solution, showing the consistent performance. The receiver clock-bias error shows some large error when the number of SVs drops to 3 and 1. However, thanks to the SLAM aiding, the bias error is constrained adequately.

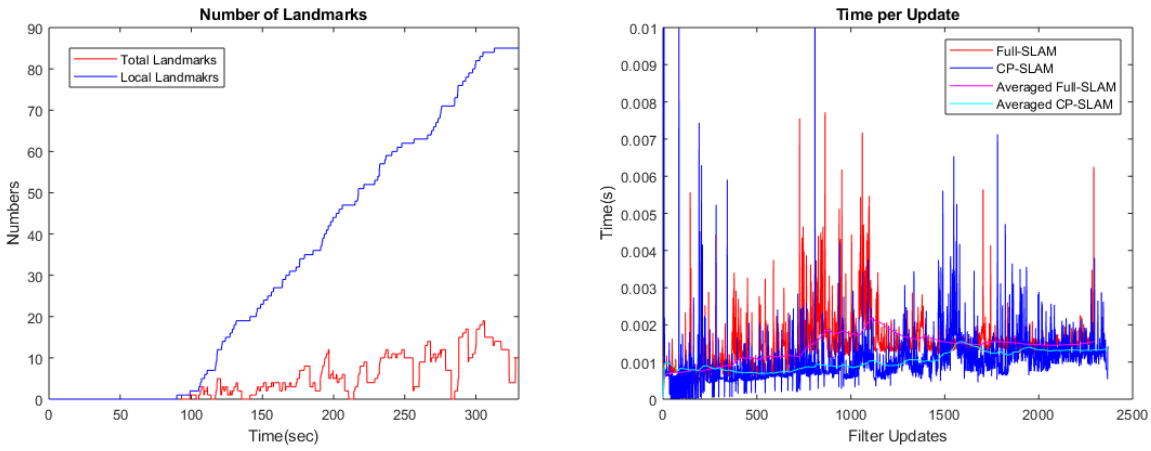
real-time processing showing on average 1.5 milliseconds processing time (using Matlab in Intel i5-core CPU with 1.7 GHz).

5 Conclusions

This article presented a pseudorange integrated SLAM (Pseudo-SLAM) which integrates GNSS pseudorange/pseudorange rates and inertial-visual SLAM to enhance the inertial-aiding during sparse landmark measurements. A computationally, efficient generalised compressed filter is implemented to achieve real-time computation. A flight dataset recorded from a high-speed, fixed-wing UAV platform was used to demonstrate the proposed method. The results showed sustained and reliable navigation solutions under a single SV and sparse landmark observations while calibrating the INS/IMU errors, and the receiver clock errors. The compressed implementation improved the computational complexity achieving the filter update time of less than 1.5 ms on average for the state dimension of 272, consisting of 17 for the vehicle states and 255 for the map (or 85 landmarks), demonstrating real-time feasibility of the method. The performance can be further improved by selecting informative satellites and landmarks, and a more precise receiver clock modelling.

Acknowledgment

Acknowledgement to ARC Discovery Project Grant number DP200101640 and Research Council of Norway grant numbers 223254 and 250725.



(a) (b)

Figure 3: (a) The comparison of the total number of landmarks registered (in blue) and the number of local landmarks in Pseudo-SLAM (in red), and (b) the comparison of the update time of the Full-SLAM (in red) and Pseudo-SLAM (in blue). The smoothed average values of the update times clearly show the benefits of the Pseudo-SLAM.

References

[Kim et al.(2017)] Kim, J.; Cheng, J.; Guivant, J.; Nieto, J.:Compressed fusion of gnss and inertial navigation with simultaneous localisation and mapping, in *IEEE Aerospace and Electronic Systems Magazine*, vol. 32, no. 8, pp. 22–36, 2017.

[Parkinson et al.(1997)] Parkinson, B.: Origins, evolution, and future of satellite navigation, in *Journal of Guidance, Control, and Dynamics*, vol. 20, no. 1, pp. 11–25, 1997.

[Skulstad et al.(2015)] Skulstad, R.; Syversen, C.; Merz, M.; Sokolova, N.; Fossen, T.; Johansen, T.: Autonomous net recovery of fixed- wing uav with single-frequency carrier-phase differential gnss," in *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 5, pp. 18–27, May 2015.

[Kim and Sukkarieh(2007)] Kim J.; Sukkarieh, S.: Real-time implementation of airborne inertial-SLAM, in *Robotics and Autonomous Systems*, vol. 55, no. 1, pp. 62–71, 2007.

[Nützi et al.(2011)] Nützi, G.;Weiss, S.; Scaramuzza, D.; Siegwart, R.: Fusion of imu and vision for absolute scale estimation in monocular slam, in *Journal of intelligent & robotic systems*, vol. 61, no. 1-4, pp. 287–299, 2011.

[Li and Mourikis(2013)] Li, M. ; Mourikis, A. I.: High-precision, consistent ekf-based visual-inertial odometry, in *The International Journal of Robotics Research*, vol. 32, no. 6, pp. 690–711, 2013.

[Sjanic et al.(2017)] Sjanic, Z.; Skoglund, M. A.; Gustafsson, F.: Em-slam with inertial/visual applications, in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 1, pp. 273–285, Feb 2017.

[Vidal et al.(2018)] Vidal, A. R.; Rebecq, H.; Horstschaefer, T.; Scaramuzza, D.: Ultimate slam? combining events, images, and imu for robust visual slam in hdr and high-speed scenarios, in *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 994–1001, 2018.

[Williams and Crump(2012)] Williams, P., and Crump, M., "All-source navigation for enhancing UAV operations in GPS-denied environments," in *Proceedings of the 28th International Congress of the Aeronautical Sciences.*, 2012.

[Guivant(2017)] Guivant, J. E.: The generalised compressed kalman filter, in *Robotica*, vol. 35, no. 8, p. 1639–1669, 2017.