

Article

Optimization Algorithm for Kalman Filter Exploiting the Numerical Characteristics of SINS/GPS Integrated Navigation Systems

Shaoxing Hu ^{1,*}, Shike Xu ¹, Duhu Wang ¹ and Aiwu Zhang ²

¹ School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China; E-Mails: xushike24k@163.com (S.X.); buaaben@163.com (D.W.)

² Ministry of Education Key Laboratory of 3D Information Acquisition and Application, Capital Normal University, Beijing 100089, China; E-Mail: zhang_aiwu@126.com

* Author to whom correspondence should be addressed; E-Mail: husx@buaa.edu.cn; Tel.: +86-10-8233-9130.

Academic Editor: Stefano Mariani

Received: 26 August 2015 / Accepted: 3 November 2015 / Published: 11 November 2015

Abstract: Aiming at addressing the problem of high computational cost of the traditional Kalman filter in SINS/GPS, a practical optimization algorithm with offline-derivation and parallel processing methods based on the numerical characteristics of the system is presented in this paper. The algorithm exploits the sparseness and/or symmetry of matrices to simplify the computational procedure. Thus plenty of invalid operations can be avoided by offline derivation using a block matrix technique. For enhanced efficiency, a new parallel computational mechanism is established by subdividing and restructuring calculation processes after analyzing the extracted “useful” data. As a result, the algorithm saves about 90% of the CPU processing time and 66% of the memory usage needed in a classical Kalman filter. Meanwhile, the method as a numerical approach needs no precise-loss transformation/approximation of system modules and the accuracy suffers little in comparison with the filter before computational optimization. Furthermore, since no complicated matrix theories are needed, the algorithm can be easily transplanted into other modified filters as a secondary optimization method to achieve further efficiency.

Keywords: computational optimization; SINS/GPS; closed-loop Kalman filter; block matrix; offline-derivation; parallel processing; accuracy-lossless decoupling; symbol operation

1. Introduction

In many SINS/GPS integrated navigation systems, common system models cannot conform to the use conditions of classical Kalman filter due to the “colored noise”. To still be able to use the optimal estimation method, it is necessary to expand the system state model [1,2]. However, as the state dimensions increase, filtering computation will rapidly become expensive and unstable, and the so-called “dimension disaster” could even break out. In order to address this problem, scholars have reported various optimization algorithms which can lower the computational costs and/or enhance the numerical robustness. The typical optimization algorithms contain square-root information filtering SRIF [3], U-D decomposition filtering [4], singular value decomposition filtering [5] and their improved versions [6–8]. These modified filters, besides greatly reducing the time consumption, theoretically guarantee the positive definiteness of covariance and effectively avoid the numerical divergence, thus they are widely applied in the theoretical design of higher-order systems. Nevertheless, they require the support of relatively complex matrix theory in derivation and may pose certain difficulty in engineering. On the other hand, although these algorithms have decreased computational costs, their complexity is still $O(n^3)$. With further expansion of the state dimensions, the above decomposition algorithms, as a kind of general method, are also being challenged on their real-time capability. Recently, researchers have paid much attention to non-commonality methods, which are designed for some specific applications or based on specific models. For instance, aiming at integrated navigation applications, a kind of reduced-order Kalman filter (RDKF) [9–12] was promoted to ease the computational load. The idea of these filters is to reduce the model dimension by theoretical/engineering methods. As the filter dimension n is vital to computation time, the reduction of the state order will produce a direct benefit in terms of real time. However, the decrease of the state order will also bring partial accuracy damage, therefore, these methods may not be suitable for some high accuracy-demanding applications. Another kind of effective algorithm optimizes the float-point operations mainly by taking advantage of the sparse matrices (e.g., the transition matrix Φ_n in SINS/GPS) during the filtering computation [13–15]. In these algorithms, the so-called matrix accumulative method (MAM) or other online methods are used and the algorithm complexity can be simplified to $O(s^2 - u^2)$ (s/u represents the numbers of nonzero/1 elements in Φ_n). As the matrices in the high order model normally contain substantial numbers of zero elements, the computational time can be curtailed to an ideal level by usage of this sparse-matrix-based method. Meanwhile, being different from RDKF, the sparse-matrix-based methods do not change the system model and need little engineering approximation, thus they perform better in the accuracy-control aspect. But these methods also have their own limitations:

- (1) Though they avoid massive unnecessary multiplying-zero operations, the methods still need extra $O(n^3)$ times estimation of zero elements in each matrix multiplication;
- (2) Methods based on MAM, a kind of online method, could not seek any deeper optimization online and their efficiency is totally dependent on the number of zeros in the real-time matrix. To enhance the optimization efficiency, the usual way is to set more zero elements by approximation methods, but this comes at the cost of a certain accuracy loss.

In response to the above problems, we present here a new highly-efficient, accuracy-lossless, and engineering-tractable optimization algorithm based on offline derivation and a parallel method

exploiting the numerical characteristics of SINS/GPS. In comparison with other algorithms, the proposed method offers numerous advantages: (i) in comparison with the general algorithms, our proposed method requires little complex deduction and is easily understandable; (ii) in comparison with the RDKF optimization method, the proposed method needs little engineering approximation and would be more accurate; and (iii) in comparison with the filter based on MAM, it is free of zero-estimation operations and can deduce some stronger conclusions, thus it is more efficient. Emphasizing on the engineering simplicity, we will introduce our optimization method revolving around the classical Kalman filter and a loosely-coupled model of SINS/GPS. It is appraised here that the chosen model does not imply any constraints on the application scope of our method. Actually, the offline and parallel method, as a pure numerical optimization method, is equally applicable to the extended filter and other complex SINS/GPS models if some similar research on the special model is done. On the other hand, since there is no transformation on any models or equations in our method (the two distinct differences with the normal filtering computations are: (i) unnecessary floating-point operations are directly avoided; and (ii) de-coupled operations are processed in parallel), the precision and robustness of our method would be equal to the classical Kalman filter in theory. Owing to this fact, the efficiency rather than robustness or accuracy is focused on in our discussion later. Section 2 gives a brief account of closed-loop Kalman filtering and a widely-used loosely-coupled model in SINS/GPS. In Section 3, the details of offline derivation and the parallel method based on a block-matrix technique are described. Section 4 and Section 5 deal with the statistics on computational costs and the filter performance on error estimation.

2. Loosely-Coupled SINS/GPS Model Using the Classical Kalman Filter

Depending on the different ways of integration, SINS/GPS are classified into two modes: (i) loose mode; and (ii) tight mode. Loose mode using GPS output to adjust SINS errors is opted for in this research for its distinctive features of high redundancy and easy realization. The emphasis is placed on a common kind of loose model *i.e.*, position-velocity integration with 18 states and six observations.

2.1. State Equation

By using the indirect method, which treats the estimation variables as parametric errors instead of parameters themselves, the state equation is established in Equation (1):

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{G}(t)\boldsymbol{\omega}(t) \quad (1)$$

where, $\mathbf{x}(t)$ represents the state vector, including attitude errors $\boldsymbol{\gamma}$, velocity errors $\delta\mathbf{v}$, position errors $(\delta L \ \delta\lambda \ \delta h)^T$, and the “colored noise” errors: gyro constant drift $\boldsymbol{\varepsilon}_c$, Markov process drift $\boldsymbol{\varepsilon}_r$, and accelerometer drift ∇_a ; $\boldsymbol{\omega}(t)$ denotes the collected system noise vector, including gyro drift white noise $\boldsymbol{\omega}_g$, Markov driving white noise $\boldsymbol{\omega}_r$, and accelerometer drift white noise $\boldsymbol{\omega}_a$; $\mathbf{A}(t)$ is state transition matrix and $\mathbf{G}(t)$ is noise coefficient matrix. Combining references [16–22] with some modification, the variables of the state equation can be described by Equations (2)–(5):

$$\mathbf{x}(t) = (\gamma_x \ \gamma_y \ \gamma_z : \delta v_x \ \delta v_y \ \delta v_z : \delta L \ \delta\lambda \ \delta h : \varepsilon_{cx} \ \varepsilon_{cy} \ \varepsilon_{cz} : \varepsilon_{rx} \ \varepsilon_{ry} \ \varepsilon_{rz} : \nabla_{ax} \ \nabla_{ay} \ \nabla_{az})^T \quad (2)$$

$$\boldsymbol{\omega}(t) = (\omega_{gx} \ \omega_{gy} \ \omega_{gz} : \omega_{rx} \ \omega_{ry} \ \omega_{rz} : \omega_{ax} \ \omega_{ay} \ \omega_{az})^T \quad (3)$$

$$A(t) = \begin{bmatrix} A_{11} & A_{12} & A_{13} & C_B^N & C_B^N & O_{3 \times 3} \\ A_{21} & A_{22} & A_{23} & O_{3 \times 3} & O_{3 \times 3} & C_B^N \\ O_{3 \times 3} & A_{32} & A_{33} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & A_{IMU22} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & A_{IMU33} \end{bmatrix}_{18 \times 18} \quad (4)$$

$$G(t) = \begin{bmatrix} C_B^N & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & I_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & I_{3 \times 3} \end{bmatrix}_{18 \times 9} \quad (5)$$

In matrices (4) and (5), all blocks are 3-order square matrices. Blocks $O_{3 \times 3}$, $I_{3 \times 3}$, and C_B^N stand for the zero matrix, the identity matrix, and the attitude matrix, respectively. The other blocks are described in detail in [16–22]. In particular, we highlight here two kinds of nonzero blocks $A_{i3(i=1,2,3)}$ and $A_{IMUii(i=2,3)}$ of which the special single-nonzero-vector and diagonal structure may contribute to deeper optimization:

$$A_{i3} = \begin{bmatrix} * & 0 & 0 \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix}, \quad i=1,2,3 \quad A_{IMUii} = \begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{bmatrix}, \quad i=2,3 \quad (6)$$

In addition, the variance $Q(t)$ of noise vector $\omega(t)$ is described as:

$$Q(t) = \begin{bmatrix} Q_{11} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & Q_{22} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & Q_{33} \end{bmatrix}_{9 \times 9} \quad (7)$$

where, block $Q_{ii(i=1,2,3)}$ is also diagonal.

2.2. Measurement Equation

We adopt a position-velocity integration model and define the filter observation vector $z(t)$ [16,22]:

$$z_{obs}(t) = \begin{bmatrix} (L_{INS} - L_{GPS})R_n \\ (\lambda_{INS} - \lambda_{GPS})R_e \cos L \\ h_{INS} - h_{GPS} \\ v_{xINS} - v_{xGPS} \\ v_{yINS} - v_{yGPS} \\ v_{zINS} - v_{zGPS} \end{bmatrix} = \begin{bmatrix} R_n \delta L + N_x \\ R_e \cos L \delta \lambda + N_y \\ \delta h + N_z \\ \delta v_x + M_x \\ \delta v_y + M_y \\ \delta v_z + M_z \end{bmatrix} \quad (8)$$

Then, we define the measurement equation as follows:

$$z(t) \triangleq H(t)x(t) + n(t) \quad (9)$$

where $n(t)$ represents the measurement noise vector; $R(t)$ denotes the corresponding variance matrix, and $H(t)$ is the coefficient matrix. $n(t)$, $R(t)$ and $H(t)$ are described as follows according to [16–22]:

$$n(t) = (N_x, N_y, N_z, M_x, M_y, M_z)^T \quad (10)$$

$$\mathbf{R}(t) = \text{diag}\{\sigma_{px}^2, \sigma_{py}^2, \sigma_{pz}^2, \sigma_{vx}^2, \sigma_{vy}^2, \sigma_{vz}^2\} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{O}_{3 \times 3} \\ \mathbf{O}_{3 \times 3} & \mathbf{R}_{22} \end{bmatrix} \quad (11)$$

$$\mathbf{H}(t) = \begin{bmatrix} \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{H}_{13} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} \\ \mathbf{O}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} \end{bmatrix}_{6 \times 18} \quad (12)$$

where:

$$\mathbf{H}_{13} = \text{diag}\{R_n, R_e \cos L, 1\} \quad (13)$$

2.3. Closed-Loop Kalman Filtering

In order to fulfill the requirement of closed-loop Kalman filtering, the state equation (Equation (1)) and measurement equation (Equation (9)) are changed to the discrete forms:

$$\mathbf{x}_n = \boldsymbol{\phi}_n \mathbf{x}_{n-1} + \mathbf{u}_{cn} + \boldsymbol{\Gamma}_n \boldsymbol{\omega}_n \quad (14)$$

$$\mathbf{z}_n = \mathbf{H}_n \mathbf{x}_n + \mathbf{n}_n \quad (15)$$

where, the transition matrix $\boldsymbol{\phi}_n$ is converted from $\mathbf{A}(t)$; the noise driving matrix $\boldsymbol{\Gamma}_n$ is converted from $\mathbf{G}(t)$; the coefficient $\mathbf{H}_n$ is converted from $\mathbf{H}(t)$; and the control vector $\mathbf{u}_{cn} = (\gamma_c^T : \delta \mathbf{v}_c^T : \delta L_c : \delta \lambda_c : \delta h_c : \boldsymbol{\varepsilon}_{cc}^T : \boldsymbol{\varepsilon}_{rc}^T : \nabla_{ac}^T)_c^T$ is added for real-time calibration on the parameters of SINS.

After discretization, the update of $\mathbf{x}_n$ is done recursively as follows:

$$\begin{aligned} \tilde{\mathbf{x}}_n(-) &= \boldsymbol{\phi}_n \tilde{\mathbf{x}}_{n-1}(+_c) \\ \tilde{\mathbf{z}}_n &= \mathbf{H}_n \tilde{\mathbf{x}}_n(-) \\ \tilde{\mathbf{x}}_n(+_e) &= \tilde{\mathbf{x}}_n(-) + \mathbf{K}_n (\mathbf{z}_{obsn} - \tilde{\mathbf{z}}_n) \\ \tilde{\mathbf{x}}_n(+_c) &= \tilde{\mathbf{x}}_n(+_e) + \mathbf{u}_{cn} \end{aligned} \quad (16)$$

where, “~” represents the parameter needed estimation; $(-)$ represents the value before estimation; $(+_e)$ represents the value after estimation; $(+_c)$ represents the value after calibration with the vector $\mathbf{u}_{cn}$; $\mathbf{z}_{obsn}$ represents real observation vector; and $\mathbf{K}_n$ represents Kalman filtering gain.

To minimize the filtering error, state vector $\tilde{\mathbf{x}}_n(+_c)$ is permanently set to zero:

$$\tilde{\mathbf{x}}_n(+_c) \equiv \mathbf{0} \quad (17)$$

Substitution of Equation (17) into Equation (16), yields:

$$\mathbf{u}_{cn} = -\tilde{\mathbf{x}}_n(+_e) = -\mathbf{K}_n \mathbf{z}_{obsn} \quad (18)$$

Apparently, $\mathbf{u}_{cn}$ becomes the final output of the Kalman filter with control vector. As $\mathbf{z}_{obsn}$ is almost directly obtained by measurement, $\mathbf{K}_n$ appears to be the key of $\mathbf{u}_{cn}$ and even the filter. According to the Kalman filtering theory, $\mathbf{K}_n$ can be recursively computed as follows:

$$\mathbf{P}_n(-) = \boldsymbol{\phi}_n \mathbf{P}_{n-1}(+_e) \boldsymbol{\phi}_n^T + \boldsymbol{\Gamma}_n \mathbf{Q}_n \boldsymbol{\Gamma}_n^T \quad (19)$$

$$\mathbf{K}_n = \mathbf{P}_n(-) \mathbf{H}_n^T (\mathbf{H}_n \mathbf{P}_n(-) \mathbf{H}_n^T + \mathbf{R}_n)^{-1} \quad (20)$$

$$\mathbf{P}_n(+_e) = (\mathbf{I} - \mathbf{K}_n \mathbf{H}_n) \mathbf{P}_n(-) \quad (21)$$

where $\mathbf{P}_n$ is the covariance; $\mathbf{Q}_n$ and $\mathbf{R}_n$ are the discretized forms of $\mathbf{Q}(t)$ and $\mathbf{R}(t)$, respectively. In fact, Equations (18)–(21) constitute the main process of the closed-loop Kalman filter. This process is

generally subdivided into two processes: time propagation Equation (19) and measurement updating Equations (18), (20) and (21). In summary, the process flow of Kalman filter in SINS/GPS is depicted in Figure 1. Since most matrices involved in Figure 1 or Equations (19)–(21) are either 18×18 or 18×6 matrices, there will be a large-scale and time-consuming computation in each recursive cycle. Therefore, a certain optimization on the computation would be essential to the real-time properties of SINS/GPS.

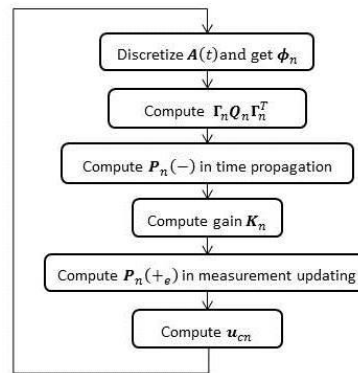


Figure 1. Iterative closed-loop Kalman filtering process.

3. Computational Optimization

Consistent with Figure 1, optimization are sequentially implemented in the updating processes of ϕ_n , $\Gamma_n Q_n \Gamma_n^T$, $P_n(-)$, K_n , $P_n(+)$, where the most computational time are consumed.

3.1. Compute ϕ_n

According to [16–22], we can calculate ϕ_n as follows:

$$\phi_n = e^{\int_{t_{n-1}}^{t_n} A(t) dt} \approx I + T_n A_n + \frac{T_n^2}{2!} A_n^2 + \frac{T_n^3}{3!} A_n^3 \quad (22)$$

Here, the 3rd order Taloy equation is selected for mainly two reasons:

- (1) Equation (22) meets the requirements of most mid-high-accuracy systems;
- (2) Equation (22) can deduce the universal optimization conclusion.

It's not difficult to prove that ϕ_n would have a uniform numerical structure independent from which of the 3rd and higher order equation is chosen. On the other hand, lower order equations are obviously a special case of 3rd order equation. Therefore, the conclusion based on 3rd equation is suitable for the model based on the different order of equations too.

As the computation of Equation (22) is centered on the 18×18 matrix $A_n = A(t_n)$, it is necessary to inspect the numerical characteristics of A_n before computational optimization: matrix A_n , in which only 64 of the total 324 elements are non-zero, is clearly a sparse matrix. A substantial reduction of useless multiplying-zero operations can be achieved by directly expanding A_n and operating every element in deduction (namely “direct derivation”), but because of the high order, direct derivation is quite complicated to handle and is not advisable in practice. Fortunately, the 3-order-block form of A_n , having 23 zero blocks among the total of 36 blocks (see Equation (4)), is still a sparse matrix structurally.

Therefore, we can carry out our offline optimization using an indirect method, in which the blocks instead of the elements are treated as the atomic unit in deduction. In this way, we can largely decrease the derivation complexity and retain most of the efficiency of a direct method at the same time. For these reasons, the indirect method is applied to all the derivation processes in this paper. In addition, the nonzero blocks A_{i3} and A_{IMUii} (see Equation (6)), which need few operations in multiplication by any 3-order block (no more than 3^2 float-pointing multiplications and 2×3 additions), are also introduced to the derivation for higher efficiency.

It is a fact that manual derivation with indirect method is still a cumbersome job in view of multiple matrix multiplications (e.g., A_n^2 , A_n^3), so we handle our derivation by a computer program instead. With the powerful symbolic-computation function of some well-known math software, e.g., MATLAB and Maple, the complicated derivation can be easily accomplished. After necessary programming, the final derivation results are shown in Equations (23) and (24):

$$A_n^2 = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} \\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} & S_{26} \\ S_{31} & S_{32} & S_{33} & O_{3 \times 1} & O_{3 \times 3} & S_{36} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & S_{55} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & S_{66} \end{bmatrix}_{18 \times 18} \quad (23)$$

$$A_n^3 = \begin{bmatrix} T_{11} & T_{12} & T_{13} & T_{14} & T_{15} & T_{16} \\ T_{21} & T_{22} & T_{23} & T_{24} & T_{25} & T_{26} \\ T_{31} & T_{32} & T_{33} & T_{34} & T_{35} & T_{36} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & T_{55} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & T_{66} \end{bmatrix}_{18 \times 18} \quad (24)$$

Also, the computation of blocks S_{ij} and T_{ij} is described in a computerized sequence as follows:

$$\begin{aligned} S_{ij} &= \sum_{k=1}^3 A_{ik} A_{kj}, \quad i, j = 1, 2, 3 \\ S_{ij} &= A_{i(j/2-1)} C_B^N, \quad i = 1, 2, 3; j = 4, 6 \\ S_{i5} &= S_{i4}, \quad i, j = 1, 2 \\ S_{15} &= S_{15} + (C_B^N A_{IMU22})_{Sim} \\ S_{26} &= S_{26} + (C_B^N A_{IMU33})_{Sim} \\ S_{55} &= (A_{IMU22}^2)_{Sim} \\ S_{66} &= (A_{IMU33}^2)_{Sim} \\ T_{ij} &= \sum_{k=1}^3 A_{ik} S_{kj}, \quad i = 1, 2, 3; j = 1, 2, \dots, 6 \\ T_{26} &= T_{26} + (C_B^N A_{IMU33}^2)_{Sim} \\ T_{55} &= (A_{IMU22} S_{55})_{Sim} \\ T_{66} &= (A_{IMU33} S_{66})_{Sim} \end{aligned} \quad (25)$$

whereas all products involving block A_{31} , S_{34} or S_{35} are zero blocks $O_{3 \times 3}$ and need no real-time computations; products involving the special blocks, e.g., A_{i3} , S_{i3} (single-nonzero –vector block), A_{i3i} , S_{55} , or S_{66} (diagonal blocks), are simplified and are partly indicated in symbol $(\cdot)_{Sim}$.

Substituting of A_n Equation (4), A_n^2 Equation (23), and A_n^3 Equation (24) in ϕ_n Equation (22), yields:

$$\phi_n = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14} & \phi_{15} & \phi_{16} \\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} & \phi_{25} & \phi_{26} \\ \phi_{31} & \phi_{32} & \phi_{33} & \phi_{34} & \phi_{35} & \phi_{36} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & I_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & \phi_{55} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & O_{3 \times 3} & \phi_{66} \end{bmatrix}_{18 \times 18} \quad (26)$$

whereas:

$$\begin{aligned} \phi_{ij} &= T_n A_{ij} + \frac{T_n^2}{2!} S_{ij} + \frac{T_n^3}{3!} T_{ij}, \quad i \in [1, 3], j \in [1, 6], i \neq j \\ \phi_{ii} &= I_{3 \times 3} + T_n A_{ii} + \frac{T_n^2}{2!} S_{ii} + \frac{T_n^3}{3!} T_{ii}, \quad i = 1, 2, 3, 5, 6 \end{aligned} \quad (27)$$

Here, ϕ_{13} and ϕ_{23} are single-nonzero-column blocks; ϕ_{33} is the sum of $I_{3 \times 3}$ and single-nonzero-column block; ϕ_{55} and ϕ_{66} are diagonal blocks. These blocks would be specially used in other processes.

In conclusion, with offline derivation based on 3rd order block method, nearly half amount of blocks ϕ_{ij} is proved to be constant blocks: $O_{3 \times 3}$ or $I_{3 \times 3}$ (see Equation (26)), thus, the computations on them are simply eliminated. Additionally, by using the properties of special blocks A_{i3} , S_{i3} , T_{i3} , A_{ij} , S_{ij} , T_{ij} and introducing the intermediate variables A_n^2 , A_n^3 , real-time computations on the other half can be further reduced.

3.2. Compute $\Gamma_n Q_n \Gamma_n^T$

According to [16–22], the 2nd order approximate of $\Gamma_n Q_n \Gamma_n^T$ are computed as follows:

$$\Gamma_n Q_n \Gamma_n^T = \frac{T_n}{2} (Q_{1n} + \phi_n Q_{1n} \phi_n^T) \quad (28)$$

where:

$$Q_{1n} = G(t)Q(t)G(t)^T = \begin{bmatrix} C_B^N Q_{11} (C_B^N)^T & O_{3 \times 9} & O_{3 \times 3} & O_{3 \times 3} \\ O_{9 \times 3} & O_{9 \times 9} & O_{9 \times 3} & O_{9 \times 3} \\ O_{3 \times 3} & O_{3 \times 9} & Q_{22} & O_{3 \times 3} \\ O_{3 \times 3} & O_{3 \times 9} & O_{3 \times 3} & Q_{33} \end{bmatrix}_{18 \times 18} \quad (29)$$

Obviously, Q_{1n} and $\Gamma_n Q_n \Gamma_n^T$ are both symmetric. Thus, for matrix $\Gamma_n Q_n \Gamma_n^T$, only its upper (or lower) triangle elements need to be computed in real time. Substitution of Q_{1n} (Equation (29)) in $\Gamma_n Q_n \Gamma_n^T$ (Equation (28)), yields:

$$\Gamma_n \mathbf{Q}_n \Gamma_n^T = \frac{T_n}{2} \begin{bmatrix} \mathbf{Q}_{n11} & \mathbf{Q}_{n12} & \mathbf{Q}_{n13} & \mathbf{O}_{3 \times 3} & \mathbf{Q}_{n15} & \mathbf{Q}_{n16} \\ \mathbf{Q}_{n12}^T & \mathbf{Q}_{n22} & \mathbf{Q}_{n23} & \mathbf{O}_{3 \times 3} & \mathbf{Q}_{n25} & \mathbf{Q}_{n26} \\ \mathbf{Q}_{n13}^T & \mathbf{Q}_{n23}^T & \mathbf{Q}_{n33} & \mathbf{O}_{3 \times 3} & \mathbf{Q}_{n35} & \mathbf{Q}_{n36} \\ \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} \\ \mathbf{Q}_{n15}^T & \mathbf{Q}_{n25}^T & \mathbf{Q}_{n35}^T & \mathbf{O}_{3 \times 3} & \mathbf{Q}_{n55} & \mathbf{O}_{3 \times 3} \\ \mathbf{Q}_{n16}^T & \mathbf{Q}_{n26}^T & \mathbf{Q}_{n36}^T & \mathbf{O}_{3 \times 3} & \mathbf{O}_{3 \times 3} & \mathbf{Q}_{n66} \end{bmatrix}_{18 \times 18} \quad (30)$$

where:

$$\begin{aligned} \mathbf{Q}_{nij} &= \frac{T_n}{2} \left[\phi_i \mathbf{Q}_{11} \phi_j^T + \phi_{i5} \left(\mathbf{Q}_{22} \phi_{j5}^T \right)_{Sim} + \phi_{i6} \left(\mathbf{Q}_{33} \phi_{j6}^T \right)_{Sim} \right], \quad i, j = 1, 2, 3; i \leq j \\ \mathbf{Q}_{n11} &= \mathbf{Q}_{n11} + \frac{T_n}{2} \mathbf{Q}_{11} \\ \mathbf{Q}_{nij} &= \frac{T_n}{2} \left(\phi_j \mathbf{Q}_{j-3, j-3} \phi_{ij} \right)_{Sim}, \quad i = 1, 2, 3; j = 5, 6 \\ \mathbf{Q}_{n55} &= \frac{T_n}{2} \left(\mathbf{Q}_{22} + \phi_{55} \mathbf{Q}_{22} \phi_{55} \right)_{Sim} \\ \mathbf{Q}_{n66} &= \frac{T_n}{2} \left(\mathbf{Q}_{33} + \phi_{66} \mathbf{Q}_{33} \phi_{66} \right)_{Sim} \end{aligned} \quad (31)$$

In Equation (31), $\mathbf{Q}_{n55}$ and $\mathbf{Q}_{n66}$ are apparently diagonal blocks that may be used to simplify other processes. As stated above, we optimize the computation of $\Gamma_n \mathbf{Q}_n \Gamma_n^T$ mainly by applying some special numerical properties to the offline derivation (e.g., the sparsity of $\mathbf{G}(t)$, the simplicity of diagonal blocks $\mathbf{Q}_{22}$, $\mathbf{Q}_{33}$, ϕ_{55} , ϕ_{66} , and the symmetry of $\Gamma_n \mathbf{Q}_n \Gamma_n^T$) and much of the computation time is saved.

3.3. Compute $\mathbf{P}_n(-)$

For simplicity, $\mathbf{P}_n(-)$ and $\mathbf{P}_n(+_e)$ will be written in a unified form $\mathbf{P}_n$ (this is allowed because of that $\mathbf{P}_n(-)$ and $\mathbf{P}_n(+_e)$ are actually the same covariance matrices in different state and share the same memories). Similar to ϕ_n or $\Gamma_n \mathbf{Q}_n \Gamma_n^T$, $\mathbf{P}_n$ is written in 3rd-order-block-matrix form:

$$\mathbf{P}_n = \begin{bmatrix} \mathbf{P}_{n,11} & \mathbf{P}_{n,12} & \cdots & \mathbf{P}_{n,16} \\ \mathbf{P}_{n,21} & \mathbf{P}_{n,22} & \cdots & \mathbf{P}_{n,26} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{n,61} & \mathbf{P}_{n,62} & \cdots & \mathbf{P}_{n,66} \end{bmatrix}_{18 \times 18} \quad (32)$$

According to its definition, $\mathbf{P}_n$ is a symmetric matrix:

$$\mathbf{P}_{n,ij} = \mathbf{P}_{n,ji}^T, \quad i, j = 1, 2, \dots, 6 \quad (33)$$

Thus, only the lower (or upper) triangle part needs substantive computation. Regardless of $\Gamma_n \mathbf{Q}_n \Gamma_n^T$ at first, Equation (19) is written as:

$$\mathbf{P}_n(-) = \phi_n \mathbf{P}_{n-1} (+_e) \phi_n^T \quad (34)$$

Substitution of Equation (26) and (32), yields:

$$\begin{aligned}
\mathbf{P}_{n,11} &= \sum_{i=1}^6 \left\{ \phi_{1i} \left[\sum_{j=1}^6 (\mathbf{P}_{n-1,ij} \phi_{1j}^T) \right]_{Temp} \right\} \\
\mathbf{P}_{n,21} &= \sum_{i=1}^6 \left\{ \phi_{2i} \left[\sum_{j=1}^6 (\mathbf{P}_{n-1,ij} \phi_{1j}^T) \right]_{Temp} \right\} \\
&\vdots \\
\mathbf{P}_{n,61} &= \left[\phi_{66} \sum_{j=1}^6 (\mathbf{P}_{n-1,6j} \phi_{1j}^T)_{Rep} \right]_{Sim} \\
\mathbf{P}_{n,22} &= \sum_{i=1}^6 \left\{ \phi_{2i} \left[\sum_{j=1}^6 (\mathbf{P}_{n-1,ij} \phi_{2j}^T) \right]_{Temp} \right\} \\
\mathbf{P}_{n,32} &= \sum_{i=1}^6 \left\{ \phi_{3i} \left[\sum_{j=1}^6 (\mathbf{P}_{n-1,ij} \phi_{2j}^T) \right]_{Temp} \right\} \\
&\vdots \\
\mathbf{P}_{n,62} &= \left[\phi_{66} \sum_{j=1}^6 (\mathbf{P}_{n-1,6j} \phi_{2j}^T)_{Rep} \right]_{Sim} \\
&\vdots \\
\mathbf{P}_{n,66} &= (\phi_{66} \mathbf{P}_{n-1,66} \phi_{66})_{Sim}
\end{aligned} \tag{35}$$

where the symbol $()_{Temp}$ represents intermediate items that would be temporarily stored in memory to avoid repeated computation; the symbol $()_{Rep}$ represents the repeated item that have been computed before and are available by directly accessing $()_{Temp}$ and the symbol $()_{Sim}$ indicates that the special properties of ϕ_{i3}^T are exploited to perform further optimization.

After Equation (35) is done, $\mathbf{\Gamma}_n \mathbf{Q}_n \mathbf{\Gamma}_n^T$ should be added. Substitution of Equations (30) and (32) in Equation (19), yields:

$$\mathbf{P}_{n,ij} = \mathbf{P}_{n,ij} + \mathbf{Q}_{n,ij}, \quad i, j = 1, 2, 3, 5, 6; i \geq j \tag{36}$$

Above all, the computation of $\mathbf{P}_n(-)$ is divided into two steps: Equations (34) and (36). As the latter has few computations, optimization mainly focuses on the former. Employing the sparsity of ϕ_n , the symmetry of $\mathbf{P}_n(-)$ and the repeatability of intermediate items, the real time of calculating Equation (34) is greatly improved.

3.4. Compute $\mathbf{K}_n$

Substitution of Equations (11), (12) and (32) in the inversion part of Equation (20), yields:

$$\mathbf{H}_n \mathbf{P}_n(-) \mathbf{H}_n^T + \mathbf{R}_n = \begin{bmatrix} \mathbf{H}_{13} \mathbf{P}_{n,33} \mathbf{H}_{13} + \mathbf{R}_{11} & \mathbf{H}_{13} \mathbf{P}_{n,23}^T \\ \mathbf{P}_{n,23} \mathbf{H}_{13} & \mathbf{P}_{n,22} + \mathbf{R}_{22} \end{bmatrix}_{6 \times 6} \tag{37}$$

Since $\mathbf{H}_n \mathbf{P}_n(-) \mathbf{H}_n^T + \mathbf{R}_n$ has a low order, few computations are needed during inversion (the complexity is $O(6^3)$ that can be neglected in comparison with other procedures) and the details of inversion are not discussed accordingly. Here, we use the symbol $\mathbf{D}$ to represent the result of inversion directly:

$$\mathbf{D} = [\mathbf{H}_n \mathbf{P}_n(-) \mathbf{H}_n^T + \mathbf{R}_n]^{-1} \tag{38}$$

Matrix $\mathbf{D}$ is easily proved to be symmetric and can be written in the 3-order-block-matrix form as:

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{12}^T & \mathbf{D}_{22} \end{bmatrix}_{6 \times 6} \quad (39)$$

Then $\mathbf{K}_n$ can be computed as:

$$\mathbf{K}_n = \begin{bmatrix} \mathbf{P}_{n,12} \mathbf{D}_{12}^T + \mathbf{P}_{n,13} \mathbf{H}_{13} \mathbf{D}_{11} & \mathbf{P}_{n,12} \mathbf{D}_{22} + \mathbf{P}_{n,13} \mathbf{H}_{13} \mathbf{D}_{12} \\ \mathbf{P}_{n,22} \mathbf{D}_{12}^T + \mathbf{P}_{n,23} \mathbf{H}_{13} \mathbf{D}_{11} & \mathbf{P}_{n,22} \mathbf{D}_{22} + \mathbf{P}_{n,23} \mathbf{H}_{13} \mathbf{D}_{12} \\ \vdots & \vdots \\ \mathbf{P}_{n,62} \mathbf{D}_{12}^T + \mathbf{P}_{n,63} \mathbf{H}_{13} \mathbf{D}_{11} & \mathbf{P}_{n,62} \mathbf{D}_{22} + \mathbf{P}_{n,63} \mathbf{H}_{13} \mathbf{D}_{12} \end{bmatrix}_{18 \times 6} \quad (40)$$

Define $\mathbf{K}_n$ as:

$$\mathbf{K}_n = \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \\ \vdots & \vdots \\ \mathbf{K}_{61} & \mathbf{K}_{62} \end{bmatrix}_{18 \times 6} \quad (41)$$

then:

$$\mathbf{K}_{ij} = \mathbf{P}_{n,i2} \mathbf{D}_{j2}^T + \mathbf{P}_{n,i3} \mathbf{H}_{13} \mathbf{D}_{1j}, \quad i = 1, 2, \dots, 6; \quad j = 1, 2 \quad (42)$$

The derivation result shows that, by using the block-matrix optimization technique, only a simple process (Equation (37), (38) and (42)) is needed in $\mathbf{K}_n$ updating. Compared with the complicated Equation (20) which needs multiple 18-order matrix production, our method seems much more simple and efficient in computing.

3.5. Compute $\mathbf{P}_n(+_e)$

Transform Equation (21) into:

$$\mathbf{P}_n(+_e) = \mathbf{P}_n(-) - \mathbf{K}_n \mathbf{H}_n \mathbf{P}_n(-) \quad (43)$$

Obviously, the pressure in computation mainly comes from $\mathbf{K}_n \mathbf{H}_n \mathbf{P}_n(-)$. Fortunately, $\mathbf{K}_n \mathbf{H}_n \mathbf{P}_n(-)$ is a symmetrical matrix (this is an evident inference according to Equation (43) in the condition that $\mathbf{P}_n(+_e)$, $\mathbf{P}_n(-)$ are both symmetrical) so that only its upper (or lower) triangular part needs real-time updating. Define $\mathbf{K}_n \mathbf{H}_n \mathbf{P}_n(-)$ as the variance's increment $\Delta \mathbf{P}_n$:

$$\Delta \mathbf{P}_n = \mathbf{K}_n \mathbf{H}_n \mathbf{P}_n(-) = \begin{bmatrix} \Delta \mathbf{P}_{11} & \Delta \mathbf{P}_{12} & \cdots & \Delta \mathbf{P}_{16} \\ \Delta \mathbf{P}_{21} & \Delta \mathbf{P}_{22} & \cdots & \Delta \mathbf{P}_{26} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta \mathbf{P}_{61} & \Delta \mathbf{P}_{62} & \cdots & \Delta \mathbf{P}_{66} \end{bmatrix}_{18 \times 18} \quad (44)$$

Substitution of Equations (12), (32) and (41), yields:

$$\Delta \mathbf{P}_{ij} = \mathbf{K}_{i1} \mathbf{H}_{13} \mathbf{P}_{n,3j} + \mathbf{K}_{i2} \mathbf{P}_{n,2j}, \quad i, j = 1, 2, \dots, 6; i \leq j \quad (45)$$

then:

$$\mathbf{P}_{n,ij} = \mathbf{P}_{n,ij} - \Delta \mathbf{P}_{ij}, \quad i, j = 1, 2, \dots, 6; i \leq j \quad (46)$$

In conclusion, taking advantage of the symmetry of $\mathbf{P}_n(+_e)$ and the sparsity of $\mathbf{H}_n$, we deduct a simple form of computing $\mathbf{P}_n(+_e)$: Equations (45) and (46). Equation (46) is apparently costless as only the

matrix subtraction is involved. Equation (45) involves matrix products, but requires no computations on most blocks of P_n (except the blocks in 2nd and 3rd rows) and proves to be highly efficient.

3.6. Parallel Computation

The optimization results (see Tables 2–6) reveal that, computation on $P_n(-)$ costs nearly half of the processing time and becomes a bottleneck for further efficiency. The low efficiency mainly arises from that $P_n(-)$ becomes increasingly dense and can provide few zero blocks after initial filtering cycles. In low-accuracy applications, some researchers may force the some matrix elements to be zero for low computational costs with the engineering approximate method. However, to some accuracy-demanded applications, methods with more accuracy and efficiency are still looked for. Parallel computation is one of the feasible approaches.

Before discussing the parallel method, we perform an inspection of the whole computation processes in order to clarify the dependency of every parameter on $P_n(-)$. The numerical relationship between $P_n(-)$ and other parameters can be consulted in Table 1.

Except for $P_n(+_e)$, all parameters are either irrelevant to $P_{n,ij}$ or only relevant to 2nd and 3rd block rows (columns) of $P_n(-)$. Taking this fact into account, we can call 2nd and 3rd block rows (columns) “useful” data and treat the rest as “useless” (it is emphasized here that “useless” only refers to the requirement of most computation processes, but not that the data is really useless and needs no updates).

Furthermore, because $P_n(-)$ is a symmetrical matrix, the “useful” blocks are only 2nd, 3rd block columns in fact. If “ $\times$ ” represents “useless” blocks, $P_n(-)$ can be written as:

$$P_n = \begin{bmatrix} \times & P_{n,12} & P_{n,31} & \times & \times & \times \\ \times & P_{n,22} & P_{n,32} & \times & \times & \times \\ \times & P_{n,32} & P_{n,33} & \times & \times & \times \\ \times & P_{n,42} & P_{n,34} & \times & \times & \times \\ \times & P_{n,52} & P_{n,35} & \times & \times & \times \\ \times & P_{n,62} & P_{n,36} & \times & \times & \times \end{bmatrix}_{18 \times 18} \quad (47)$$

Table 1. Dependency of every real-time parameter on $P_n(-)$ blocks.

$\phi_n, Q_n, P_n(-)$	K_n (Equations (40) and (41))	$P_n(+_e)$ (Equation (43))	ΔP_{ij} (Equation (45))
Irrelevant to $P_{n,ij}$	Dependent on $P_{n,i2}, P_{n,i3}$ only	Dependent on $P_{n,ij}$	Dependent on $P_{n,2i}, P_{n,3i}$ only

In measurement updating (Equations (18), (20) and (21)), all parameters except $P_n(+_e)$ are decoupled with the “useless” data (Equation (19)) which belong to the time propagation. Moreover, to the coupled $P_n(+_e)$ its time-consuming part ΔP_{ij} is also independent of the “useless” blocks. In other words, with some appropriate modification of measurement updating and time propagation, we can produce a new decoupling mechanism of the two procedures. The modification contains three steps:

Step1. Subdivide time propagation.

Classify the blocks of $P_n(-)$ into two kinds: “useful” data and “useless” data;

Step 2. Subdivide measurement updating.

Subdivide $P_n(+_e)$ into two processes: computing ΔP_{ij} (Equation (45)) and adding $P_n(-)$ (Equation (46));
 Step 3. Restructure measurement updating.

Restructure measurement updating process as ΔP_{ij} (instead of $P_n(+_e)$) (Equations (18) and (21)).

With the above steps, the restructured process and the “useless” data updating process are completely numerically decoupled and can be computed in parallel. In this way, the efficiency bottleneck in $P_n(-)$ updating is broken.

Furthermore, in order to achieve the parallel computation on the “useful” data as well, we similarly subdivide the process of updating “useful” data into two steps: computing $\phi_n P_{n-1}(+_e) \phi_n^T$ (Equation (34)) and adding $\Gamma_n Q_n \Gamma_n^T$ (Equation (36)). Apparently, computing $\phi_n P_{n-1}(+_e) \phi_n^T$ and updating $\Gamma_n Q_n \Gamma_n^T$ (Equation (31), which belongs to time propagation) can be processed in parallel. In summary, the process of Kalman filtering with parallel computation is described as Figure 2.

It is pointed out here that the parallel method in this paper has a fundamental difference with some other parallel methods [23,24]. The others owe their decoupling to the so-called “mandatory delay”, which may cause certain accuracy damage. On the contrary, the decoupling in this paper needs neither “mandatory delay” nor engineering approximation and is achieved mainly by introducing the “useful” data and subdividing the computation process. Essentially, it is the numerical characteristic of SINS/GPS that brings up this special decoupling. Therefore, the parallel processing in this paper is an accuracy-lossless method.

In respect of efficiency, the parallel optimization method can reduce much execution time needed in updating both $\Gamma_n Q_n \Gamma_n^T$ and the “useless” blocks (including 918 multiplications and 603 additions in updating $\Gamma_n Q_n \Gamma_n^T$, and 1089 multiplications and 1107 additions in updating “useless” blocks). In comparison with non-parallel methods, efficiency of parallel method increases by about 25%.

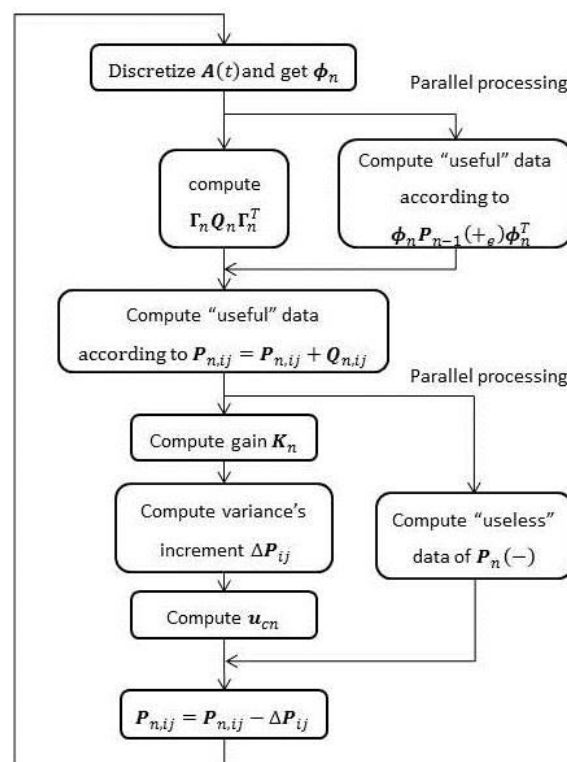


Figure 2. Filtering process with parallel computation.

4. Optimization Efficiency

According to the derivation result, we can figure out the optimization efficiency by counting and comparing the computational counts needed in both the proposed method and the existing algorithm. In detail, the discussion about counts proceeds on each parameter updating procedure, respectively. As matrix multiplication is mainly involved, the calculation will focus on the floating-point operations of multiplication and addition. Tables 2–6 will show the statistical details of each parameter. Then, Tables 7 and 8 give the global efficiency and the comparison with other general algorithms, respectively. As shown in the tables, the offline-derivation and parallel method saves about 90% of computational time and 66% of memory space, while keeping the required accuracy level. Compared with general decomposition optimization algorithms, the proposed method needs no complicated matrix theory but producing higher efficiency. Under the condition that high hardware precision has guaranteed the numerical robustness, our algorithm appears quite simple and practical.

Table 2. Optimization efficiency of ϕ_n updating process.

	Offline-Derivation Method	Traditional Method
Multiplication	$42 \times 3^3 + 36 \times 3^2 + 13 \times 3 \times 3^2 = 1809$ (I), 10%	$3 \times 18^3 + 3 \times 18^2 = 18468$ (II)
Addition	$42 \times 2 \times 3^2 + 9 \times 2 \times 3 + 13 \times 2 \times 3^2 = 1044$ (I), 6%	$3 \times 17 \times 18^2 + 3 \times 18^2 = 17496$ (II)
Memory	$32 \times 3^2 \times 8 \text{ Byte} = 2.3 \text{ KB}$ (III)	$3 \times 18^2 \times 8 \text{ Byte} = 7.6 \text{ KB}$ (III)

In (I), 42×3^3 ($42 \times 2 \times 3^2$), 36×3^2 ($9 \times 2 \times 3$) and $13 \times 3 \times 3^2$ ($13 \times 2 \times 3^2$) represent the counts on multiplication (addition) of items in Equation (25) without special blocks, of items in Equation (25) with special block and of items in Equation (27) respectively, and only 3^2 & 3^3 items are counted; In (II), the first item represents the counts on operations of 18-order-matrix multiplication, the second item represents that of multiplication by scalars or matrix addition; In (III), data are assumed to be double-precision (8-byte word length); the percentages represent the efficiency in comparison with the traditional method.

Table 3. Optimization efficiency of $\Gamma_n Q_n \Gamma_n^T$ updating process.

	Offline-Derivation Method	Traditional Method
Multiplication	$24 \times 3^3 + 18 \times 3^2 + 12 \times 3^2 = 918$, 6%	$2 \times 18^3 + 18^2 + 18 \times 6^2 + 6 \times 18^2 = 14580$
Addition	$24 \times 2 \times 3^2 + 9 \times 2 \times 3 + 13 \times 3^2 = 603$, 4%	$2 \times 17 \times 18^2 + 18^2 + 5 \times 6 \times 18 + 17 \times 18 \times 6 = 13716$
Memory	$13 \times 3^2 \times 8 \text{ Byte} = 0.9 \text{ KB}$	$3 \times 18^2 \times 8 \text{ Byte} = 7.6 \text{ KB}$

Table 4. Optimization efficiency of $P_n(-)$ updating process.

	Offline-Derivation Method	Traditional Method
Multiplication	$120 \times 3^3 + 37 \times 3^2 = 3573$, 31%	$2 \times 18^3 = 11664$
Addition	$120 \times 2 \times 3^2 + 18 \times 2 \times 3 + 134 \times 3^2 = 3474$, 31%	$2 \times 17 \times 18^2 + 18^2 = 11340$
Memory	$60 \times 3^2 \times 8 \text{ Byte} = 4.2 \text{ KB}$	$3 \times 18^2 \times 8 \text{ Byte} = 7.6 \text{ KB}$

Table 5. Optimization efficiency of κ_n updating process.

	Offline-Derivation Method	Traditional Method
Multiplication	$24 \times 3^3 + 6 \times 3^2 = 702$, 11%	$18 \times 18 \times 6 + 18 \times 6 \times 6 + 6 \times 18 \times 18 + 6 \times 18 \times 6 = 5184$
Addition	$24 \times 2 \times 3^2 + 12 \times 3^2 = 540$, 11%	$17 \times 18 \times 6 + 17 \times 6 \times 6 + 5 \times 18 \times 6 + 17 \times 18 \times 6 = 4824$
Memory	$16 \times 3^2 \times 8 \text{ Byte} = 1.1 \text{ KB}$	$(3 \times 6 \times 18 + 6 \times 6) \times 8 \text{ Byte} = 2.8 \text{ KB}$

Table 6. Optimization efficiency of $P_n(+_e)$ updating process.

	Offline-Derivation Method	Traditional Method
Multiplication	$42 \times 3^3 + 6 \times 3^2 = 1188$, 15%	$6 \times 18 \times 18 + 18^3 = 7776$
Addition	$42 \times 2 \times 3^2 + 42 \times 3^2 = 1134$, 15%	$5 \times 18^2 + 17 \times 18^2 + 18^2 = 7452$
Memory	$16 \times 3^2 \times 8 \text{ Byte} = 1.1 \text{ KB}$	$(3 \times 6 \times 18 + 6 \times 6) \times 8 \text{ Byte} = 2.8 \text{ KB}$

Table 7. Optimization efficiency of offline-derivation and parallel method.

	Offline-Derivation & Parallel Method	Traditional Method
Multiplication	$1809 + 2484 + 702 + 1188 = 6183$ (IV), 10.7%	$18468 + 14580 + 11664 + 5184 + 7776 = 57672$
Addition	$1044 + 2475 + 540 + 1134 = 5193$ (IV), 9.5%	$17496 + 13716 + 11340 + 4824 + 7452 = 54828$
Memory	$2.3 + 0.9 + 4.2 + 1.1 + 1.1 = 9.6 \text{ KB}$	$7.6 + 7.6 + 7.6 + 2.8 + 2.8 = 28.4 \text{ KB}$

In (IV), 2484 (2475) is the operation times of “useful” blocks updating.

Table 8. Comparison of offline-derivation & parallel method and general optimization filters.

	Multiplication	Addition
SRIF filtering	$\frac{7}{6} \times 18^3 + 36 \times 18^2 = 18468$ (V)	$\frac{7}{6} \times 18^3 + 36 \times 18^2 = 18468$ (V)
U-D decomposing filtering	$\frac{1}{2} \times 18^3 + \frac{1}{2} \times 18^2 + 10 \times 18^2 + 8 \times 18 = 6462$ (VI)	$\frac{1}{2} \times 18^3 + \frac{1}{2} \times 18^2 + 9 \times 18^2 + 9 \times 18 = 6156$ (VI)
SVD filtering	$26 \times 18^3 + 78 \times 18^2 = 176904$ (VII)	$26 \times 18^3 + 78 \times 18^2 = 176904$ (VII)
offline-derivation and parallel method	$2484 + 702 + 1188 = 4374$ (VIII)	$2475 + 540 + 1134 = 4149$ (VIII)

(V) is explained in Reference [3], while (VI) in Reference [25], Reference [26] and (VII) in Reference [5]; Point out here that, (VI) relies on the precondition that ϕ_n is already an upper-triangle matrix in 9-order-block form (what is the conclusion of offline derivation), otherwise, 9702 multiplications and 9396 additions are actually needed; concerning (VIII), the procedure of computing ϕ_n is no longer counted as no optimizations are against ϕ_n in the general algorithms.

5. Simulation

To evaluate the method validity, we design a practical loosely-coupled SINS/GPS program in CCS v5.5 (a well-known programming tool on DSP platform). The program comprises three modules:

- (1) Sensor data sampling module.
- (2) SINS algorithm module. This module is designed to implement the calculation of attitude, velocity and position independently (only the inertial sensors data is needed).
- (3) Kalman filter module. This part is for data fusion. The system module is designed according to Section 2, and the computation process is programmed using the offline and parallel method

described in Section 3. Besides, for accuracy evaluation, the classical KF is also designed as a control group.

With the developed program, we can simulate filtering process and evaluate the performance of our method in CCS. The simulation parameters and conditions are set as follows:

- (1) DSP TMS320C6722, a famous float-point CPU, is chosen in CCS as the computing device.
- (2) The carrier of the navigation system is assumed to be static. With this assumption, the gyro/accelerometer would detect no valid rotate/velocity rate aside from devices noise. Therefore, noise is considered the sampling data driving the SINS/GPS program. As the carrier is static, velocity/position should stay at the zero/initial value in theory, thus, this assumption can largely facilitate the evaluation of program results.
- (3) Velocity/position measured by GPS is constantly set:

$$v_{GPS} = (0, 0, 0)^T, (L \ \lambda \ h)_{GPS}^T = (0.6977686 \ 2.0306152 \ 90.0)^T$$

Velocity/position calculated by SINS is initialized:

$$v_{SINS} = (0, 0, 0)^T, (L \ \lambda \ h)_{SINS}^T = (0.6977688370723981 \ 2.030615370778482 \ 91.0)^T$$

- (4) IMU data rate (noise frequency): 100 Hz; SINS velocity/position updating rate: 50/20 Hz; filtering rate: 10 Hz.

With the above setting, we run the program in CCS for 100 s and record the navigation outputs (e.g., the X-axis velocity and latitude). The simulation results are illustrated in Figures 3 and 4.

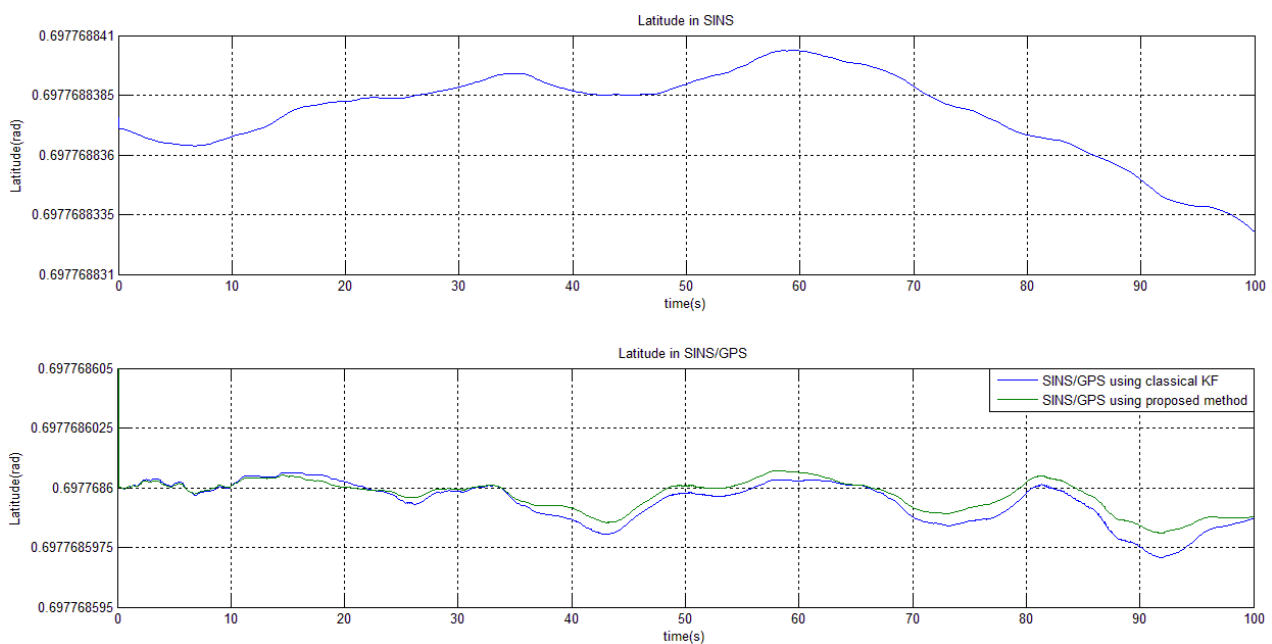


Figure 3. Latitude calculation by different methods.

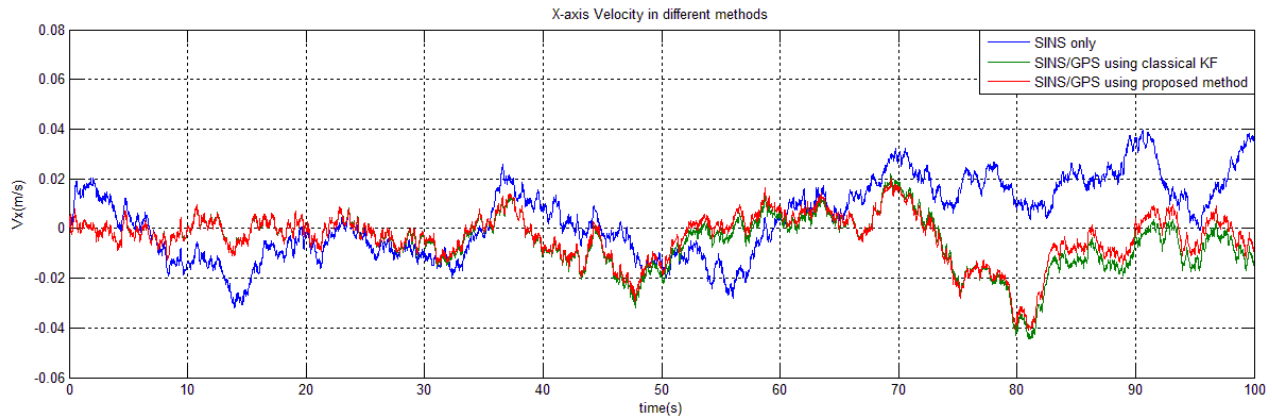


Figure 4. Velocity calculation by different methods.

The results show that the proposed method effectively restrains the unbounded error accumulation in SINS and keep the estimated variables around their mean values. On the aspect of estimation performance, our method is nearly at par to the classical KF (the proposed-method curve does not completely overlap the classical-KF one because of less calculation and less truncation errors). This conclusion agrees with the theoretical analysis: our numerical method, needing no modification of the system module and no engineering approximation, causes little damage to the estimation accuracy, but from the point of view of real-time computation, our method is much more efficient than the classical KF. In the DSP program, the proposed method needs 64,820 system clocks in each filtering cycle while classical KF needs 390,480 clocks. If a clock frequency of 200 MHz, the typical frequency of a DSP TMS320C6722 is chosen, then our method will cost only 324 μ s in each filtering operation. Such a level of real-time efficiency is highly valuable to the many accurate navigation applications.

6. Conclusions

Starting from a common classical-Kalman-fitter-based SINS/GPS model (position-speed integration with 18 states and six measurements), we present an optimization algorithm based on the system's numerical characteristics. The algorithm employs a block-matrix technique in offline derivation, where special blocks (e.g., zero blocks, diagonal blocks, *etc.*) are used to simplify the calculation. In this way, plenty of invalid multiplying by zero and repeated operations are avoided offline. Furthermore, aiming at the time-consuming update of $P_n(-)$, a novel parallel method is implemented by defining "useful" data and subdividing computational process, and the entire filtering procedure is greatly accelerated. The statistical analysis and simulation results reveal that the offline algorithm, coupled with the parallel method, can reduce the CPU processing time by 90% and the memory usage by 66%. Compared with several general decomposition-optimization algorithms, the proposed method requires no complex matrix theory rather more efficiency. The distinguished feature of the algorithm is that no modification of the system module and no engineering approximation are needed, thus it causes little harm to the base filter. It is pointed out that although the derivation is based on a specific integration model, the algorithm as a complete numerical optimization approach can be transplanted to other advanced models of SINS/GPS or other extended filters. With the powerful symbolic-operation function of the MATLAB program, researchers can manipulate formulas involving high order matrices in an easy way.

Consequently, the proposed method is an engineering-tractable approach with high efficiency and high precision for SINS/GPS integrated navigation systems.

Author Contributions

Shaoxing Hu and Shike Xu developed the overall algorithm and wrote the papers. Shike Xu coded the program and performed the simulation. Wangdu Hu analyzed the experiment data. Aiwu Zhang reviewed and revised the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Mohinder, S.G.; Angus, P.A. *Global Navigation Satellite Systems, Inertial Navigation, and Integration*; John Wiley & Sons Inc.: New York, NY, USA, 2013.
2. Chui, C.K.; Chen, G. *Kalman Filtering with Real-Time Applications*; Tsinghua University Press: Beijing, China, 2013.
3. Bierman, G.J. A comparison of discrete linear filtering algorithms. *IEEE Trans. Aero. Electr. Syst.* **1973**, *9*, 28–37.
4. Thornton, C.L.; Bierman, G.J. UDUT covariance factorization for Kalman filtering. *Control Dyn.* **1980**, *16*, 177–248.
5. Wang, L.; Libert, G.; Manneback, P. Kalman filter algorithm based on singular value decomposition. In Proceedings of the 31st IEEE Conference on Decision and Control, Tucson, AZ, USA, 16–18 December 1992.
6. Boncelet, C.; Dickinson, B. An extension of the SRIF Kalman filter. *IEEE Trans. Autom. Control* **1987**, *32*, 176–179.
7. Tsyganova, J.V.; Kulikova, M.V. State sensitivity evaluation within UD based array covariance filters. *IEEE Trans. Autom. Control* **2013**, *58*, 2944–2950.
8. Tang, Y.; Deng, Z.; Manoj, K.K.; Chen, D. A practical scheme of the sigmapoint Kalman filter for high-dimensional systems. *J. Adv. Model. Earth Syst.* **2014**, *6*, 21–37.
9. He, X.; Chen, Y.; Iz, H.B. A reduced-order model for integrated GPS/INS. *IEEE Aerosp. Electron. Syst. Mag.* **1998**, *13*, 40–45.
10. Huang, J.; Tan, H.S. A low-order DGPS-based vehicle positioning system under urban environment. *IEEE ASME Trans. Mechatron.* **2006**, *11*, 567–575.
11. Papazoglou, M.; Tsioras, C. Integrated SAR/GPS/INS for Target Geolocation Improvement. *J. Comput. Model.* **2014**, *4*, 267–298.
12. Mutambara, A.G. *Decentralized Estimation and Control for Multisensor Systems*; CRC Press: Boca Raton, FL, USA, 1998.
13. Zhu, Q.J.; Yan, G.M. A Rapid Computation Method for Kalman Filtering in Vehicular SINS/GPS Integrated System. *Appl. Mech. Mater.* **2012**, *182*, 541–545.

14. Wei, C.; Fang, J.; Sheng, J. Fast data fusion method for integrated navigation system and hardware in loop simulation. *J. Beijing Univ. Aeronaut. Astronaut.* **2006**, *11*, 1281–1285. (In Chinese)
15. Holmes, S.; Klein, G.; Murray, D.W. An $O(N^3)$ Square Root Unscented Kalman Filter for Visual Simultaneous Localization and Mapping. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 1251–1263.
16. González, R.; Giribet, J.I.; Patiño, H.D. An approach to benchmarking of loosely coupled low-cost navigation systems. *Math. Comput. Model. Dyn. Syst.* **2015**, *21*, 272–287.
17. Quinchia, A.G.; Falco, G.; Falletti, E.; Dosis, F.; Ferrer, C. A comparison between different error modeling of MEMS applied to GPS/INS integrated systems. *Sensors* **2013**, *13*, 9549–9588.
18. Falco, G.; Einicke, G.A.; Malos, J.T.; Dosis, F. Performance analysis of constrained loosely coupled GPS/INS integration solutions. *Sensors* **2012**, *12*, 15983–16007.
19. Solimeno, A. Low-cost INS/GPS Data Fusion with Extended Kalman Filter for Airborne Applications. Master's Thesis, Universidade Technica de Lisboa, Lisboa, Portugal, 2007.
20. Farrell, J.A.; Barth, M. *The Global Positioning Systems and Inertial Navigation*; McGraw-Hill: New York, NY, USA, 1999.
21. Wang, H. *GPS Navigation Principle and Application*; Science Press: Beijing, China, 2010. (In Chinese)
22. Wu, T.; Ma, L. *Strapdown Inertial Navigation System Application Analysis*; National Defence Industry Press: Beijing, China, 2011. (In Chinese)
23. Brown, D.W.; Gaston, F.M.F. The design of parallel square-root covariance Kalman filters using algorithm engineering. *Integr. VLSI J.* **1995**, *20*, 101–119.
24. Rosén, O.; Medvedev, A. Parallelization of the Kalman filter on multicore computational platforms. *Control Eng. Pract.* **2013**, *21*, 1188–1194.
25. Bierman, G.J. Measurement Updating Using the U-D Factorization. *Automatica* **1976**, *12*, 375–382.
26. Bierman, G.J. Efficient Time Propagation of U-D Covariance Factors. *IEEE Trans. Autom. Control* **1981**, *26*, 890–894.