

Article

A Novel Real-Time Autonomous Localization Algorithm Based on Weighted Loosely Coupled Visual–Inertial Data of the Velocity Layer

Cheng Liu ¹, Tao Wang ^{1,2} , Zhi Li ^{1,2,*}  and Peng Tian ^{3,*}

¹ State Key Laboratory of Explosion Science and Safety Protection, Beijing Institute of Technology, Beijing 100081, China; 3220235030@bit.edu.cn (C.L.); wang_tao@bit.edu.cn (T.W.)

² Beijing Institute of Technology Chongqing Innovation Center, Chongqing 401100, China

³ School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 401100, China

* Correspondence: lizhi_haligong@163.com (Z.L.); tianpeng@cqjtu.edu.cn (P.T.)

Abstract: IMUs (inertial measurement units) and cameras are widely utilized and combined to autonomously measure the motion states of mobile robots. This paper presents a loosely coupled algorithm for autonomous localization, the ICEKF (IMU-aided camera extended Kalman filter), for the weighted data fusion of the IMU and visual measurement. The algorithm fuses motion information on the velocity layer, thereby mitigating the excessive accumulation of IMU errors caused by direct subtraction on the positional layer after quadratic integration. Furthermore, by incorporating a weighting mechanism, the algorithm allows for a flexible adjustment of the emphasis placed on IMU data versus visual information, which augments the robustness and adaptability of autonomous motion estimation for robots. The simulation and dataset experiments demonstrate that the ICEKF can provide reliable estimates for robot motion trajectories.

Keywords: loosely coupled visual–inertial measurement; visual inertial odometry; robot autonomous localization; data fusion



Academic Editor: Alessandro Di Nuovo

Received: 10 December 2024

Revised: 14 January 2025

Accepted: 17 January 2025

Published: 20 January 2025

Citation: Liu, C.; Wang, T.; Li, Z.; Tian, P. A Novel Real-Time Autonomous Localization Algorithm Based on Weighted Loosely Coupled Visual–Inertial Data of the Velocity Layer. *Appl. Sci.* **2025**, *15*, 989. <https://doi.org/10.3390/app15020989>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Motivation

As the demand for autonomous mobility of robots continues to increase, localization technology continues to evolve. When robots perform motion tasks in environments such as extraterrestrial planets, rugged ground terrains, aerial spaces, or maritime environments [1–3], they face various challenges, including unstable satellite signals, impact and vibration upon sensors, limited visual features, and unpredictable light conditions. Furthermore, autonomous localization is a requisite ability for aerial [4], ground [5,6], or maritime [3,7] autonomous robots to successfully accomplish challenging tasks of local planning aiming for navigation to prevent collisions.

Integrated visual–inertial sensors do not require active signal emission or depend on external preset references, which benefits further data fusion as autonomous odometry. Two primary fusion strategies exist to facilitate fusion: loosely coupled methods and tightly coupled methods [8].

Loosely coupled methods are computationally efficient and maintain a higher update rate than tightly coupled methods do, thereby enabling low-cost and compact implementations. This approach additionally allows the simultaneous integration of localization information from multiple sources such as visual measurement, GNSS (global navigation

satellite system), LIDAR (light detection and ranging), etc. [9,10]. Especially for application on the multiagent [11], loosely coupled methods are naturally capable of processing simultaneous multi-source information. Although the estimation accuracy is relatively low owing to the lack of an optimization process, which tight-coupled methods [12,13] usually perform, the integration of multiple measurement data into a loosely coupled framework is inevitable to address the challenges of environmental uncertainty during practical utilization. Therefore, a loosely coupled data fusion approach that is capable of adjusting the weights of the measurement results from individual sensors or agents is necessary to achieve a balanced output, considering how much the information source can be trusted.

1.2. Related Work

Loosely coupled visual–inertial odometry methods are usually constructed based on filter frameworks such as the EKF [14,15]. High-confidence visual keyframe measurements are then periodically utilized to correct the high-frequency integration results from IMUs. This basic structure effectively balances the updating rate and accuracy.

Kelly and Sukhatme [16] proposed a data fusion algorithm for the self-calibration of a monocular visual–inertial system with proven observability. With a similar filtering structure, Weiss and Siegwart [17] incorporated the world coordinate system drift in visual measurements, as well as the spatial relationships between the IMU and the monocular camera, into the EKF framework. This approach enables failure detection and scale-drift estimation by fusing of measurement data on the position and attitude layers.

Achtelik and Weiss [18] constructed a filter-based framework to recover the relative configuration of two drones performing IMUs and monocular visual measurements. To address error amplification during the error propagation process in EKFs due to possible inaccurate state modeling, BROSSARD et al. [19] proposed the visual–inertial invariant EKF algorithm, which is based on Lie algebra rules. Furthermore, researchers designed visual–inertial odometry that embraces the characteristics of novel filters such as the MEKF [20] and the equivariant filter [21,22] to broaden the application of filter-based techniques in motion estimation.

The aforementioned loosely coupled strategies have demonstrated remarkable performance when mainly focused on the fusion of positional layers. However, since the high-frequency update and the intrinsic drift of IMUs inevitably result in fast error accumulation, the effect of the final correction on the position and attitude layers will be suboptimal when the visual measurement update frequency drops or accuracy deteriorates. Furthermore, these methods are thus incapable of adjusting the emphasis on either source while encountering external visual interference.

1.3. Our Approach

This paper presents a loosely coupled algorithm for autonomous localization with a weighted data fusion on the linear velocity and angular velocity layers, namely the ICEKF. The weight of the IMU measurement can be adjusted according to whether the external conditions are favorable for visual measurement, thus increasing the overall accuracy of autonomous localization. The real-time performance is promising since the time complexity is $O(n)$. The local weak observability of the ICEKF is demonstrated.

The organization of this paper is as follows: the state vector of the ICEKF is established in the second section, the design of the propagation and update of the filter is discussed in the third section, the observability analysis is presented in the fourth section, and the fifth section carries out numerical studies of simulations and dataset experiments. Finally, a summary is provided.

2. Design of the State Vector

This section discusses the construction of the state vector and analyzes the weighted coupling process on the velocity layer. The forms of the ICEKF state vector are then deduced.

2.1. Definition of Variables in the ICEKF

The descriptions of the main coordinate frames and variables used in this study are listed in Table 1. The relationships between the coordinate frames of the IMU-aided camera integration system are established, as shown in Figure 1, where p_c^i and \bar{q}_c^i represent the relative linear translation and rotation between the IMU and the camera, which are considered constants once calibrated. The IMU outputs linear acceleration a_{mi} and angular velocity ω_{mi} in its rigid body coordinate frame. This paper considers the visual black-box measurement outputting rotation \bar{q}_w^c and unscaled linear translation p_w^c in the world frame. The fusion result is attached to the IMU coordinate frame, with the coupled linear translation and rotation denoted as p_w^{ic} and \bar{q}_w^{ic} , respectively. The derivatives of the translation and rotation are v_w^{ic} and $\dot{\bar{q}}_w^{ic}$. The units used in this paper are described in the SI, such as acceleration in m^2/s and angular velocity in rad/s .

Table 1. Coordinate frames and notations in the ICEKF.

Symbol	Description
w	fixed world coordinate frame
i	coordinate frame attached to the IMU
c	coordinate frame attached to the camera
ic	coordinate frame attached to the IMU-aided camera system
x_B^A	x represents a general viable vector, A is the coordinate frame attached to the vector, and B is the reference frame; for example, p_w^c denotes the linear translation of the camera with the frame c , measured with respect to the world frame w
p	translation vector of rigid bodies along 3 axes, of which the quasi-quaternion description is $\bar{p} = [0, p^T]^T$
\bar{q}	unit quaternion according to the Hamilton notation [16], written as $\bar{q} = [q_0, q_1, q_2, q_3]^T = [q_0, \mathbf{q}^T]^T$
\bar{q}^*	conjugate quaternion of \bar{q} , and $\bar{q}^* \otimes \bar{q} = 1$
R	rotation matrix converted from \bar{q} , such as $R_w^i = R(\bar{q}_w^i)$
$[x \times]$	skew-symmetric matrix of x , and $[x \times]y = -x[y \times]$ [23]
n	white Gaussian noise vector with zero mean and covariance σ^2
g	gravity vector in the world frame

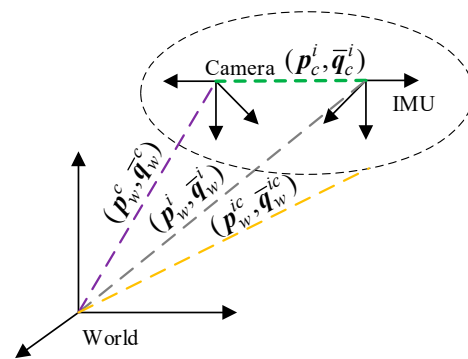


Figure 1. Relationships of the coordinate frames in the ICEKF.

2.2. Construction of the State Vector

Given the necessity for information fusion on the velocity layer, while containing first-order derivatives and integrals, the state of the ICEKF is defined as a column vector consisting of 30 elements, as follows:

$$\mathbf{X} = \left\{ \mathbf{p}_w^{icT}, \mathbf{v}_w^cT, \mathbf{v}_w^iT, \bar{\mathbf{q}}_w^{icT}, \bar{\mathbf{q}}_w^iT, \boldsymbol{\omega}_c^cT, \boldsymbol{\omega}_i^iT, \mathbf{b}_{ai}T, \mathbf{b}_{\omega i}T, \lambda \right\}^T, \tag{1}$$

where \mathbf{v}_w^c is the linear velocity of the visual measurement derived from \mathbf{p}_w^c ; $\boldsymbol{\omega}_c^c$ is the equivalent body angular velocity of the camera measured in the IMU frame; $\bar{\mathbf{q}}_w^i$ and $\boldsymbol{\omega}_i^i$ are the rotation velocity and body angular velocity, respectively; \mathbf{b}_{ai} and $\mathbf{b}_{\omega i}$ are the biases of \mathbf{a}_{mi} and $\boldsymbol{\omega}_{mi}$, respectively; λ is the scale coefficient of the monocular visual translation. The remaining variables are described in the last section. Let \mathbf{a}_w^i be the linear acceleration of the IMU measured in the world frame, where $\mathbf{a}_w^i = \mathbf{R}_w^i(\mathbf{a}_{mi} - \mathbf{b}_{ai} - \mathbf{n}_{a_{mi}})$. Additionally, $\boldsymbol{\omega}_i^i = \boldsymbol{\omega}_{mi} - \mathbf{b}_{\omega i} - \mathbf{n}_{\omega_{mi}}$.

The relationships among the variables in the ICEKF state are shown in Figure 2.

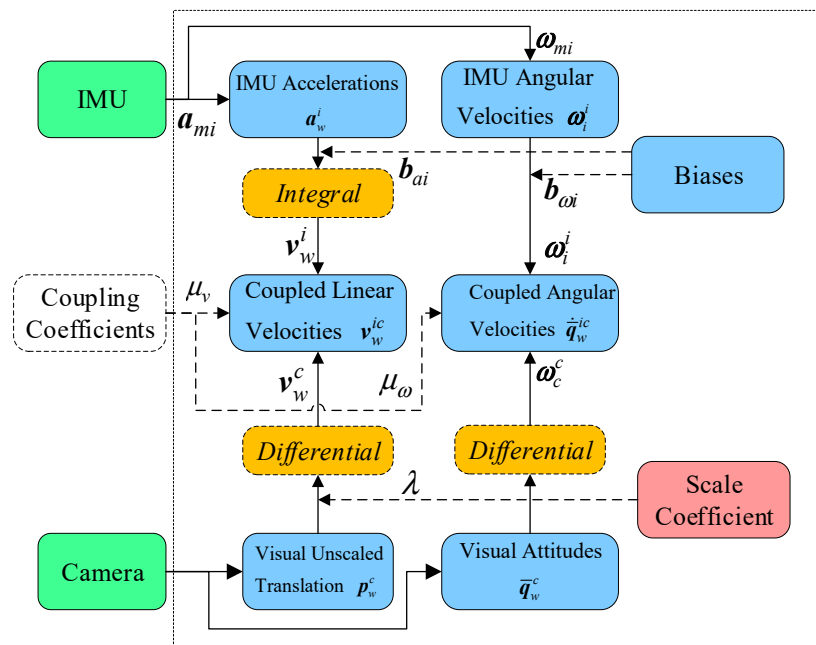


Figure 2. Relationships among the variables in the ICEKF state vector.

When the visual measurement is treated as a black box, $\mathbf{v}_w^c = \dot{\mathbf{p}}_w^c$. To obtain $\boldsymbol{\omega}_c^c, \bar{\mathbf{q}}_w^c$ should be transformed to $\bar{\mathbf{q}}_w^{c'}$, which is the visual rotation described in the IMU frame. Then, $\boldsymbol{\omega}_c^c$ can be obtained from the derivatives of the Euler angles recovered from $\bar{\mathbf{q}}_w^{c'}$. The coupling coefficients of the linear and angular velocities, $\mu_v, \mu_\omega \in [0, 1]$, are independent of the state update. This design guarantees that the coupling weights are always adjustable, which increases the accuracy of motion estimation without impacting observability.

To focus on the filter construction in this paper, the following is assumed:

1. The drift of the visual world frame in the visual black-box measurement is negligible due to its slow variation.
2. The outcomes acquired through the first-order derivation of measurement results are modeled as random walk, including $\boldsymbol{\omega}_c^c$ and \mathbf{v}_w^c .

2.3. Coupling Process

The coupling process of the linear velocities of the IMU-camera system requires a weighted summation as follows:

$$\mathbf{v}_w^{ic} = \mu_v \mathbf{v}_w^i + \lambda(1 - \mu_v) \mathbf{R}_c^i \mathbf{v}_w^c, \tag{2}$$

The angular velocity, which covaries with the rotation of the coordinate frames, cannot be directly summed. According to Appendix A and the proposed lemmas in Appendix B, with $\mu_\omega \in [0, 1]$, the following equations hold:

$$\frac{d}{dt} (\bar{\mathbf{q}}_w^i)^{\mu_\omega} = \frac{1}{2} (\bar{\mathbf{q}}_w^i)^{\mu_\omega} \otimes (\mu_\omega \bar{\boldsymbol{\omega}}_i^i), \tag{3}$$

$$\frac{d}{dt} (\bar{\mathbf{q}}_w^{c'})^{1-\mu_\omega} = \frac{1}{2} (\bar{\mathbf{q}}_w^{c'})^{1-\mu_\omega} \otimes ((1 - \mu_\omega) \bar{\boldsymbol{\omega}}_c^c), \tag{4}$$

where $\bar{\mathbf{q}}_w^{c'}$ is the equivalent rotation of the visual measurement in the IMU frame.

After further deduction according to Appendix A, the derivative of $\bar{\mathbf{q}}_w^{ic}$ is written as follows:

$$\begin{aligned} \dot{\bar{\mathbf{q}}}_w^{ic} &= \frac{d}{dt} (\bar{\mathbf{q}}_w^i)^{\mu_\omega} \otimes (\bar{\mathbf{q}}_w^{c'})^{1-\mu_\omega} + (\bar{\mathbf{q}}_w^i)^{\mu_\omega} \otimes \frac{d}{dt} (\bar{\mathbf{q}}_w^{c'})^{1-\mu_\omega} \\ &= \frac{\mu_\omega}{2} (\bar{\mathbf{q}}_w^i)^{\mu_\omega} \otimes \bar{\boldsymbol{\omega}}_i^i \otimes (\bar{\mathbf{q}}_w^{c'})^{1-\mu_\omega} + \frac{1-\mu_\omega}{2} \bar{\mathbf{q}}_w^i \otimes \bar{\boldsymbol{\omega}}_c^c \end{aligned}, \tag{5}$$

where $\bar{\boldsymbol{\omega}} = [0, \boldsymbol{\omega}]$ is the supplementary vector of the body angular vector $\boldsymbol{\omega}$.

Angular velocity fusion according to Equation (5) is performed while assuming that $\bar{\mathbf{q}}_w^{ic} = \bar{\mathbf{q}}_w^{c'} = \bar{\mathbf{q}}_w^i$, which are all attached to the IMU frame. Thus, potential errors can be caused by the assembly and measurement noise for a real IMU-camera integrated rig. However, after careful calibration and the filter process, the influence of the possible inconsistency is negligible.

2.4. Simplification of the State Vector

To simplify further discussion, the subscripts of the variables described in the world coordinate system are omitted; thus, the ICEKF state vector in Equation (1) is rewritten as follows:

$$\mathbf{X} = \left\{ \mathbf{p}_{ic}^T \ \mathbf{v}_c^T \ \mathbf{v}_i^T \ \bar{\mathbf{q}}_{ic}^T \ \bar{\mathbf{q}}_i^T \ \boldsymbol{\omega}_c^T \ \boldsymbol{\omega}_i^T \ \mathbf{b}_a^T \ \mathbf{b}_w^T \ \lambda \right\}^T, \tag{6}$$

The derivatives of the variables in the vector are as follows:

$$\begin{aligned} \dot{\mathbf{p}}_{ic} &= \mathbf{v}_{ic} = \mu_v \mathbf{v}_i + \lambda(1 - \mu_v) \mathbf{R}_c^i \mathbf{v}_c, \ \dot{\mathbf{v}}_c = \mathbf{n}_{v_c}, \ \dot{\mathbf{v}}_i = \mathbf{R}_i (\mathbf{a}_{mi} - \mathbf{b}_{ai} - \mathbf{n}_{a_{mi}}) - \mathbf{g}, \\ \dot{\bar{\mathbf{q}}}_{ic} &= \frac{\mu_\omega}{2} \bar{\mathbf{q}}_i^{\mu_\omega} \otimes \bar{\boldsymbol{\omega}}_i \otimes \bar{\mathbf{q}}_c^{1-\mu_\omega} + \frac{1-\mu_\omega}{2} \bar{\mathbf{q}}_{ic} \otimes \bar{\boldsymbol{\omega}}_c, \ \dot{\bar{\mathbf{q}}}_i = \frac{1}{2} \bar{\mathbf{q}}_i \otimes \bar{\boldsymbol{\omega}}_i, \ \dot{\boldsymbol{\omega}}_c = \mathbf{n}_{\boldsymbol{\omega}_c}, \ \dot{\boldsymbol{\omega}}_i = \frac{d}{dt} (\boldsymbol{\omega}_{mi} - \mathbf{b}_{\boldsymbol{\omega}_i} - \mathbf{n}_{\boldsymbol{\omega}_{mi}}), \\ \dot{\mathbf{b}}_{ai} &= \mathbf{n}_{\mathbf{b}_{ai}}, \ \dot{\mathbf{b}}_{wi} = \mathbf{n}_{\mathbf{b}_{wi}}, \ \dot{\lambda} = \mathbf{n}_\lambda, \end{aligned} \tag{7}$$

where \mathbf{v}_c , $\boldsymbol{\omega}_c$, and λ are modeled as random walk, and notably $\bar{\mathbf{q}}_c = \bar{\mathbf{q}}_c^i \otimes \bar{\mathbf{q}}_w^c = \bar{\mathbf{q}}_w^{c'}$.

2.5. Error of the State Vector

Let $\hat{\mathbf{X}}$ denote the expectation of \mathbf{X} , and let $\tilde{\mathbf{X}}$ denote the error between the expectation, and the measurement, written as $\tilde{\mathbf{X}} = \mathbf{X} - \hat{\mathbf{X}}$. $\tilde{\mathbf{X}}$ with 28 elements, is described as follows:

$$\tilde{\mathbf{X}} = \left\{ \Delta \mathbf{p}_{ic}^T \ \Delta \mathbf{v}_c^T \ \Delta \mathbf{v}_i^T \ \delta \boldsymbol{\theta}_{ic}^T \ \delta \boldsymbol{\theta}_i^T \ \Delta \boldsymbol{\omega}_c^T \ \Delta \boldsymbol{\omega}_i^T \ \Delta \mathbf{b}_{ai}^T \ \Delta \mathbf{b}_{wi}^T \ \Delta \lambda \right\}^T, \tag{8}$$

With the small-angle assumption, when the rotation angle corresponding to a unit quaternion $\bar{\mathbf{q}}$ is very small, the error of $\bar{\mathbf{q}}$ is written as $\delta \bar{\mathbf{q}} = [q_0, \delta \mathbf{q}^T]^T \approx \left[1, \frac{1}{2} \delta \boldsymbol{\theta}^T \right]^T$ [23]. Because the algorithm operates at a high update rate, the high-order terms that yield

negligible computational results are disregarded, such as $\delta q \cdot \Delta \omega$, $\delta q \cdot \delta q$, and $\delta q \cdot n$. Then, the description of the derivative $\dot{\tilde{X}}$ can be inspected.

According to Equation (2), the derivative of the coupled translation error is as follows:

$$\Delta \dot{p}_{ic} = \mu_v \Delta v_i + \lambda(1 - \mu_v) R_c^i \Delta v_c, \tag{9}$$

The expectation of the linear acceleration of the IMU is $\hat{a}_i = a_{mi} - \hat{b}_{ai}$, and $\Delta b_{ai} = b_{ai} - \hat{b}_{ai}$. According to Appendix A, the rotation can be reformed as $R_i = R(\bar{q}_i) \approx \hat{R}_i(I_3 + [\delta \theta_i \times])$ under the small angle assumption. After neglecting the high-order terms, the error of the linear velocity of the IMU is written as follows:

$$\begin{aligned} \Delta \dot{v}_i &= a_i - \hat{a}_i = R_i(a_{mi} - b_{ai} - n_{a_{mi}}) - g - \hat{R}_i(a_{mi} - \hat{b}_{ai}) + g, \\ &\approx -\hat{R}_i[\hat{a}_i \times] \delta \theta_i - \hat{R}_i \Delta b_{ai} - n_{a_{mi}} \end{aligned} \tag{10}$$

For general quaternions and angular velocities, $\bar{q} = \hat{q} \otimes \delta \bar{q}$ and $\bar{\omega} = \hat{\omega} + \Delta \bar{\omega}$. By subjecting $\bar{q}_{ic}^* = (\bar{q}_c)^{1-\mu_\omega} \otimes (\bar{q}_i)^{\mu_\omega}$ and $(\bar{q}_i^{\mu_\omega})^* \otimes \bar{q}_i^{\mu_\omega} = 1$ to Equation (5), the error of the coupled rotation can be deduced as follows:

$$\begin{aligned} \delta \dot{\bar{q}}_{ic} &= \hat{q}_{ic}^* \otimes \left(\frac{\mu_\omega}{2} \bar{q}_i^{\mu_\omega} \otimes \bar{\omega}_i \otimes \bar{q}_c^{1-\mu_\omega} + \frac{1-\mu_\omega}{2} \bar{q}_{ic} \otimes \bar{\omega}_c - \frac{\mu_\omega}{2} \hat{q}_i^{\mu_\omega} \otimes \hat{\omega}_i \otimes \hat{q}_c^{1-\mu_\omega} \otimes \delta \bar{q}_{ic} - \frac{1-\mu_\omega}{2} \hat{q}_{ic} \otimes \hat{\omega}_c \otimes \delta \bar{q}_{ic} \right) \\ &= \left[\frac{\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \bar{q}_{ic}^* \otimes \bar{q}_i^{\mu_\omega} \otimes (\hat{\omega}_i + \Delta \bar{\omega}_i) \otimes \bar{q}_c^{1-\mu_\omega} - \frac{\mu_\omega}{2} (\hat{q}_c^{1-\mu_\omega})^* \otimes (\hat{q}_i^{\mu_\omega})^* \otimes \hat{q}_i^{\mu_\omega} \otimes \hat{\omega}_i \otimes \hat{q}_c^{1-\mu_\omega} \otimes \delta \bar{q}_{ic} \right] \\ &+ \left[\frac{1-\mu_\omega}{2} \delta \bar{q}_{ic} \otimes (\hat{\omega}_c + \Delta \bar{\omega}_c) - \frac{1-\mu_\omega}{2} \hat{\omega}_c \otimes \delta \bar{q}_{ic} \right] \\ &= \left[\begin{array}{l} \frac{\mu_\omega}{2} \delta \bar{q}_{ic} \otimes (\bar{q}_c^{1-\mu_\omega})^* \otimes \hat{\omega}_i \otimes \bar{q}_c^{1-\mu_\omega} \\ + \frac{\mu_\omega}{2} \delta \bar{q}_{ic} \otimes (\bar{q}_c^{1-\mu_\omega})^* \otimes \Delta \bar{\omega}_i \otimes \bar{q}_c^{1-\mu_\omega} \\ - \frac{\mu_\omega}{2} (\hat{q}_c^{1-\mu_\omega})^* \otimes \hat{\omega}_i \otimes \hat{q}_c^{1-\mu_\omega} \otimes \delta \bar{q}_{ic} \end{array} \right] + \left[\begin{array}{l} \frac{1-\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \hat{\omega}_c \\ + \frac{1-\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \Delta \bar{\omega}_c \\ - \frac{1-\mu_\omega}{2} \hat{\omega}_c \otimes \delta \bar{q}_{ic} \end{array} \right] \end{aligned} \tag{11}$$

According to Appendix A, the rotation of p_{ic} with respect to $\bar{q}_c^{1-\mu_\omega}$ is written as follows:

$$(\hat{q}_c^{1-\mu_\omega})^* \otimes \bar{p}_{ic} \otimes \hat{q}_c^{1-\mu_\omega} \triangleq \hat{R}(\hat{q}_c^{1-\mu_\omega})^T p_{ic} = \hat{R}_{\mu_\omega c}^T p_{ic}, \tag{12}$$

where $R_{\mu_\omega c}^T$ is the rotation matrix converted from $\bar{q}_c^{1-\mu_\omega}$.

By subjecting Equation (12) to Equation (11) and disregarding the high-order terms, the following is obtained:

$$\begin{aligned} \delta \dot{\bar{q}}_{ic} &= \left\{ \begin{array}{l} \frac{\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \left[\begin{array}{l} 0 \\ R_{\mu_\omega c}^T \hat{\omega}_i \end{array} \right] \\ - \frac{\mu_\omega}{2} \left[\begin{array}{l} 0 \\ R_{\mu_\omega c}^T \hat{\omega}_i \end{array} \right] \otimes \delta \bar{q}_{ic} \\ + \frac{\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \left[\begin{array}{l} 0 \\ R_{\mu_\omega c}^T \Delta \omega_i \end{array} \right] \end{array} \right\} + \left[\begin{array}{l} \frac{1-\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \hat{\omega}_c \\ - \frac{1-\mu_\omega}{2} \hat{\omega}_c \otimes \delta \bar{q}_{ic} \\ + \frac{1-\mu_\omega}{2} \delta \bar{q}_{ic} \otimes \Delta \bar{\omega}_c \end{array} \right] \\ &= \frac{\mu_\omega}{2} \left\{ \begin{array}{l} \left[\begin{array}{l} 0 \\ R_c^T \hat{\omega}_i \\ R_c^T \hat{\omega}_i \end{array} \right] - \left[\begin{array}{l} 0 \\ R_{\mu_\omega c}^T \hat{\omega}_i \\ R_{\mu_\omega c}^T \hat{\omega}_i \end{array} \right] \\ + \left[\begin{array}{l} 0 \\ R_c^T \Delta \omega_i \\ R_c^T \Delta \omega_i \end{array} \right] - \left[\begin{array}{l} 0 \\ R_{\mu_\omega c}^T \Delta \omega_i \\ R_{\mu_\omega c}^T \Delta \omega_i \end{array} \right] \end{array} \right\} + \frac{1-\mu_\omega}{2} \left\{ \begin{array}{l} \left[\begin{array}{l} 0 \\ \hat{\omega}_c \\ \hat{\omega}_c \end{array} \right] - \left[\begin{array}{l} 0 \\ \hat{\omega}_c \\ \hat{\omega}_c \end{array} \right] \\ + \left[\begin{array}{l} 0 \\ \Delta \omega_c \\ \Delta \omega_c \end{array} \right] - \left[\begin{array}{l} 0 \\ \Delta \omega_c \\ \Delta \omega_c \end{array} \right] \end{array} \right\} \\ &\approx -\frac{\mu_\omega}{2} \left\{ \begin{array}{l} \left[\begin{array}{l} 0 \\ 2[(R_{\mu_\omega c}^T \hat{\omega}_i) \times] \delta q_{ic} \\ q_0 R_{\mu_\omega c}^T \Delta \omega_i \end{array} \right] - \left[\begin{array}{l} 0 \\ q_0 R_{\mu_\omega c}^T \Delta \omega_i \end{array} \right] \end{array} \right\} - \frac{1-\mu_\omega}{2} \left\{ \begin{array}{l} \left[\begin{array}{l} 0 \\ 2[\hat{\omega}_c \times] \delta q_{ic} \\ q_0 \Delta \omega_c \end{array} \right] - \left[\begin{array}{l} 0 \\ q_0 \Delta \omega_c \end{array} \right] \end{array} \right\}, \end{aligned} \tag{13}$$

Then, with the small angle assumption $\delta\bar{q}_{ic} = [q_0, \delta q_{ic}^T]^T \approx [1, \frac{1}{2}\delta\theta_{ic}^T]^T$, the error of the coupled equivalent Euler angle is simplified as follows:

$$\delta\dot{\theta}_{ic} = -\mu_{\omega} \left[\left(R_{\mu_{\omega}c}^T \hat{\omega}_i \right) \times \delta\theta_{ic} - R_{\mu_{\omega}c}^T \Delta\omega_i \right] - (1 - \mu_{\omega}) \left[\hat{\omega}_c \times \delta\theta_{ic} - \Delta\omega_c \right], \quad (14)$$

The expectation of the body angular velocity of the IMU is $\hat{\omega}_i = \omega_{mi} - \hat{b}_{\omega i}$, and $\Delta b_{\omega i} = b_{\omega i} - \hat{b}_{\omega i}$. Resembling the process above, the error of the rotation of the IMU is written as follows:

$$\begin{aligned} \delta\dot{\bar{q}}_i &= \hat{q}_i^* \otimes (\dot{\bar{q}}_i - \dot{\hat{q}}_i \otimes \delta\bar{q}_i) = \frac{1}{2}\delta\bar{q}_i \otimes (\hat{\omega}_i + \Delta\bar{\omega}_i) - \frac{1}{2}\hat{\omega}_i \otimes \delta\bar{q}_i \\ &= \frac{1}{2} \begin{bmatrix} 0 & -\hat{\omega}_i \\ \hat{\omega}_i & -[\hat{\omega}_i \times] \end{bmatrix} \begin{bmatrix} q_{i0} \\ \delta q_i \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 & -\hat{\omega}_i \\ \hat{\omega}_i & [\hat{\omega}_i \times] \end{bmatrix} \begin{bmatrix} q_{i0} \\ \delta q_i \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & -\Delta\omega_i \\ \Delta\omega_i & -[\Delta\omega_i \times] \end{bmatrix} \begin{bmatrix} q_{i0} \\ \delta q_i \end{bmatrix}, \\ &\approx \begin{bmatrix} 0 \\ -[\hat{\omega}_i \times] \delta q_i \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{2} q_{i0} \Delta\omega_i \end{bmatrix} \end{aligned} \quad (15)$$

Equation (15) can be simplified with the small angle assumption as follows:

$$\delta\dot{\theta}_i = -[\hat{\omega}_i \times] \delta\theta_i - \Delta\omega_i = -[\hat{\omega}_i \times] \delta\theta_i - \Delta b_{\omega i} - n_{\omega_{mi}}, \quad (16)$$

The derivatives of the remaining terms in \tilde{X} are as follows:

$$\dot{\Delta v}_c = n_{v_c}, \dot{\Delta\omega}_c = n_{\omega_c}, \dot{\Delta\omega}_i = \frac{d}{dt}(\Delta b_{\omega i} + n_{\omega_{mi}}) = n_{\omega_i}, \dot{\Delta b}_{ai} = n_{b_{ai}}, \dot{\Delta b}_{\omega i} = n_{b_{\omega i}}, \dot{\Delta\lambda} = n_{\lambda}, \quad (17)$$

3. Propagation and Update of the ICEKF

The propagation and update of the ICEKF are described in detail in this section. The key matrices in the propagation step determine the internal transition process of the filter. The update step rectifies the filtered outcomes in reference to the measurement results.

3.1. Propagation

For the linearized continuous-time errors of an ICEKF state, the following equation exists [24]:

$$\dot{\tilde{X}} = F_c \tilde{X} + G_c n, \quad (18)$$

where $n = [n_{\omega_c}^T, n_{a_{mi}}^T, n_{\omega_c}^T, n_{\omega_{mi}}^T, n_{b_{ai}}^T, n_{b_{\omega i}}^T]^T$ is the propagation noise vector for the ICEKF following the Gaussian distribution. F_c and G_c are considered constant during every iteration.

To discretize Equation (18) during the period Δt , we write the description of the discrete state transition matrix F_d and the discrete noise covariance matrix Q_d as follows [24]:

$$\begin{cases} F_d = \exp(F_c \Delta t) = I_d + F_c \Delta t + \frac{1}{2!} F_c^2 \Delta t^2 + \dots \\ Q_d = \int_t^{t+\Delta t} F_d(\tau) G_c Q_c G_c^T F_d(\tau)^T d\tau \end{cases}, \quad (19)$$

where $Q_c = \text{diag}(\sigma_{n_{v_c}}^2, \sigma_{n_{a_i}}^2, \sigma_{n_{\omega_c}}^2, \sigma_{n_{\omega_i}}^2, \sigma_{n_{b_{ai}}}^2, \sigma_{n_{b_{\omega i}}}^2)$ is the diagonal matrix converted from Gaussian noise.

Section 2.5 presents the specific expression of \tilde{X} . Considering that the algorithm operates at a high frequency, only the first-order expansion in Equation (19) is considered. The complete expression of F_d is derived as follows:

$$F_d = \begin{bmatrix} I_3 & A_1 & A_2 & \dots & \mathbf{0}_{3 \times 19} \\ \mathbf{0}_3 & I_3 & \dots & \mathbf{0}_{3 \times 22} \\ \mathbf{0}_3 & \mathbf{0}_3 & I_3 & \mathbf{0}_3 & A_3 & \mathbf{0}_3 & \mathbf{0}_3 & A_4 & \mathbf{0}_3 & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & A_5 & \mathbf{0}_3 & A_6 & A_7 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & A_8 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & A_9 & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{13 \times 15} & \dots & I_{13 \times 13} & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}_{28 \times 28}, \quad (20)$$

where $A_1 = \lambda(1 - \mu_v)\mathbf{R}_c^i \Delta t$, $A_2 = \mu_v \mathbf{I}_3 \Delta t$, $A_3 = -\hat{\mathbf{R}}_i [\hat{\mathbf{a}}_i \times] \Delta t$, $A_4 = -\hat{\mathbf{R}}_i \Delta t$, $A_5 = \mathbf{I}_3 - (\mu_\omega [(\mathbf{R}_{\mu_\omega c}^T \hat{\boldsymbol{\omega}}_i) \times] + (1 - \mu_\omega) [\hat{\boldsymbol{\omega}}_c \times]) \Delta t$, $A_6 = (1 - \mu_\omega) \mathbf{I}_3 \Delta t$, $A_7 = \mu_\omega \mathbf{R}_{\mu_\omega c}^T \Delta t$, $A_8 = \mathbf{I}_3 - [\hat{\boldsymbol{\omega}}_I \times] \Delta t$, and $A_9 = -\mathbf{I}_3 \Delta t$.

Q_d can be further derived by combining Equation (20) and G_c , and the explicit form of G_c is recovered according to Equation (18) as follows:

$$G_c = \begin{bmatrix} \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ I_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & -\hat{\mathbf{R}}_i & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & -I_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & I_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & I_3 & \mathbf{0}_3 \\ \mathbf{0}_{7 \times 18} & \dots & \dots & \dots & \dots & \dots \end{bmatrix}_{28 \times 18}. \quad (21)$$

3.2. Measurement

Let the state measurement matrix at the k -th step of the ICEKF be z_k . After the measured linear translation z_p , the attitude quaternion z_q , the linear velocity z_v , and the angular velocity z_ω are obtained by comparing the results from the keyframes of the visual measurement with the expectation of the ICEKF, z_k is obtained as follows:

$$\tilde{z}_k = \begin{bmatrix} \tilde{z}_p^T & \tilde{z}_q^T & \tilde{z}_v^T & \tilde{z}_\omega^T \end{bmatrix}^T, \quad (22)$$

For the error of the linear translation, the following holds:

$$\begin{aligned} \tilde{z}_p &= (\mathbf{p}_{ic} - \mathbf{R}_{ic} \mathbf{p}_i^c) \lambda + \mathbf{n}_p - (\hat{\mathbf{p}}_{ic} - \hat{\mathbf{R}}_{ic} \hat{\mathbf{p}}_i^c) \hat{\lambda} \\ &= (\mathbf{p}_{ic} \lambda - \hat{\mathbf{p}}_{ic} \hat{\lambda}) + (\hat{\mathbf{p}}_{ic} \lambda - \hat{\mathbf{p}}_{ic} \hat{\lambda}) + [\hat{\mathbf{R}}_{ic} \mathbf{p}_i^c \lambda - \hat{\mathbf{R}}_{ic} (\mathbf{I}_3 + [\delta \boldsymbol{\theta}_{ic} \times]) \mathbf{p}_i^c \lambda] + (-\hat{\mathbf{R}}_{ic} \mathbf{p}_i^c \lambda + \hat{\mathbf{R}}_{ic} \hat{\mathbf{p}}_i^c \hat{\lambda}) + \mathbf{n}_p, \\ &= \lambda \Delta \mathbf{p} + \hat{\mathbf{R}}_{ic} [(\mathbf{p}_i^c \lambda) \times] \delta \boldsymbol{\theta}_{ic} + (\hat{\mathbf{p}}_{ic} - \hat{\mathbf{R}}_{ic} \mathbf{p}_i^c) \Delta \lambda + \mathbf{n}_p \end{aligned} \quad (23)$$

For the error of the attitude quaternion considering the small angle assumption, the following equation is used:

$$\tilde{z}_q = \hat{\mathbf{q}}_{ic}^* \otimes (\hat{\mathbf{q}}_c^i \otimes \hat{\mathbf{q}}_c) = \delta \bar{\mathbf{q}}_{ic} \approx \begin{bmatrix} 1 \\ \frac{1}{2} \delta \boldsymbol{\theta}_{ic} + \mathbf{n}_\theta \end{bmatrix}, \quad (24)$$

For the error of the linear velocity, the following holds:

$$\begin{aligned} \tilde{z}_v &= \mathbf{v}_{ic} - \hat{\mathbf{v}}_{ic} \\ &= \mu_v \Delta \dot{\mathbf{v}}_i \Delta t + \lambda(1 - \mu_v) \mathbf{R}_c^i \Delta \mathbf{v}_c \\ &= -\mu_v \Delta t \hat{\mathbf{R}}_i [\hat{\mathbf{a}}_i \times] \delta \boldsymbol{\theta}_i - \mu_v \Delta t \hat{\mathbf{R}}_i \Delta \mathbf{b}_a + \lambda(1 - \mu_v) \mathbf{R}_c^i \Delta \mathbf{v}_c + \mathbf{n}_v \end{aligned}, \quad (25)$$

Since the angular velocities in different coordinate frames cannot be directly subtracted, the measurement error can be indirectly obtained according to Equation (14) as follows:

$$\begin{aligned} \tilde{z}_\omega &= \delta\dot{\theta}_{ic}\Delta t + \mathbf{n}_\omega \\ &= -\Delta t [\mu_\omega [(\mathbf{R}_{\mu_\omega c}^T \hat{\omega}_i) \times] + (1 - \mu_\omega) [\hat{\omega}_c \times]] \delta\theta_{ic} + (1 - \mu_\omega) \Delta t \Delta \omega_c + \mu_\omega \Delta t \mathbf{R}_{\mu_\omega c}^T \Delta \omega_i + \mathbf{n}_\omega \end{aligned} \quad (26)$$

Based on the description of \tilde{X} in Equation (8) and the above expansion of the measurement process, under the small angle assumption, the measurement error is reformulated as follows:

$$\tilde{z}_k \simeq \mathbf{H}_k \tilde{X}_k + \mathbf{n}_m, \quad (27)$$

where \mathbf{H}_k is the measurement matrix, and the measurement noise is simplified as $\mathbf{n}_m = [\mathbf{n}_p^T \mathbf{n}_\theta^T \mathbf{n}_v^T \mathbf{n}_\omega^T]^T$. Thus, \mathbf{H}_k can be recovered from Equations (22)–(26) as follows:

$$\mathbf{H}_k = \begin{bmatrix} \lambda \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 6} & \mathbf{B}_1 & \mathbf{0}_{3 \times 15} & \mathbf{B}_2 \\ \mathbf{0}_{3 \times 9} & \frac{1}{2} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 6} & \mathbf{0}_{3 \times 16} & \\ \mathbf{0}_{3 \times 3} & \mathbf{B}_3 & \mathbf{0}_{3 \times 6} & \mathbf{B}_4 & \mathbf{0}_{3 \times 6} & -\mu_v \Delta t \hat{\mathbf{R}}_i & \mathbf{0}_{3 \times 4} \\ \mathbf{0}_{3 \times 9} & \mathbf{B}_5 & \mathbf{0}_{3 \times 3} & (1 - \mu_\omega) \Delta t \mathbf{I}_{3 \times 3} & \mu_\omega \Delta t \mathbf{R}_{\mu_\omega c}^T & \mathbf{0}_{3 \times 7} \end{bmatrix}_{12 \times 28}, \quad (28)$$

where $\mathbf{B}_1 = \hat{\mathbf{R}}_{ic} [(\mathbf{p}_i^c \lambda) \times]$, $\mathbf{B}_2 = \hat{\mathbf{p}}_{ic} - \hat{\mathbf{R}}_{ic} \mathbf{p}_i^c$, $\mathbf{B}_3 = \lambda \Delta t (1 - \mu_v) \mathbf{R}_{ic}^i$, $\mathbf{B}_4 = -\mu_v \Delta t \hat{\mathbf{R}}_i [\hat{\mathbf{a}}_i \times]$ and $\mathbf{B}_5 = -\Delta t [\mu_\omega [(\mathbf{R}_{\mu_\omega c}^T \hat{\omega}_i) \times] + (1 - \mu_\omega) [\hat{\omega}_c \times]]$.

3.3. Entire ICEKF Process

The entire process of the ICEKF in the k -th iteration is presented as follows:

Step 1. By calculating F_d and Q_d according to Equation (19), the prior covariance matrix of errors $P_{k+1|k}$ can be obtained from the following:

$$P_{k+1|k} = F_d P_{k|k} F_d^T + Q_d, \quad (29)$$

Step 2. The Kalman gain matrix K_k is updated as follows:

$$\begin{cases} S_k = \mathbf{H}_k P_{k+1|k} \mathbf{H}_k^T + \mathbf{R}_k \\ K_k = P_{k+1|k} \mathbf{H}_k^T S_k^{-1} \end{cases}, \quad (30)$$

where \mathbf{R}_k is the measurement noise matrix.

Step 3. The current state $\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + \hat{X}_k$ is calculated according to $\hat{X}_k = K_k \tilde{z}_k$, where \tilde{z}_k is obtained via Equation (27).

Step 4. The posterior covariance matrix of errors $P_{k+1|k+1}$ is updated, and Step 1 is carried out again according to the following:

$$P_{k+1|k+1} = (\mathbf{I}_{28 \times 28} - K_k \mathbf{H}_k) P_{k+1|k} (\mathbf{I}_{28 \times 28} - K_k \mathbf{H}_k)^T + K_k \mathbf{R}_k K_k^T, \quad (31)$$

During Step 2, the measurement error model of the visual measurement can be preestablished as described in [25,26]. While updating the attitude during Step 3, the angular error of rotation in \hat{X}_k is achieved with the small angle assumption. Therefore, the quaternion form of the attitude expectation \hat{q}_{k+1} must be recovered as follows [23]:

$$\begin{cases} \delta \hat{q}_k \approx \frac{1}{2} \delta \hat{\theta}_k \\ \hat{q}_{k+1} = \begin{cases} \left[\sqrt{1 - \delta \hat{q}_k^T \delta \hat{q}_k}, \delta \hat{q}_k^T \right]^T & \text{if } \delta \hat{q}_k^T \delta \hat{q}_k \leq 1 \\ \frac{1}{\sqrt{1 + \delta \hat{q}_k^T \delta \hat{q}_k}} [1, \delta \hat{q}_k^T]^T & \text{if } \delta \hat{q}_k^T \delta \hat{q}_k > 1 \end{cases} \end{cases}, \quad (32)$$

For a clearer view of the entire process, the data flow during the k -th iteration is illustrated by Figure 3.

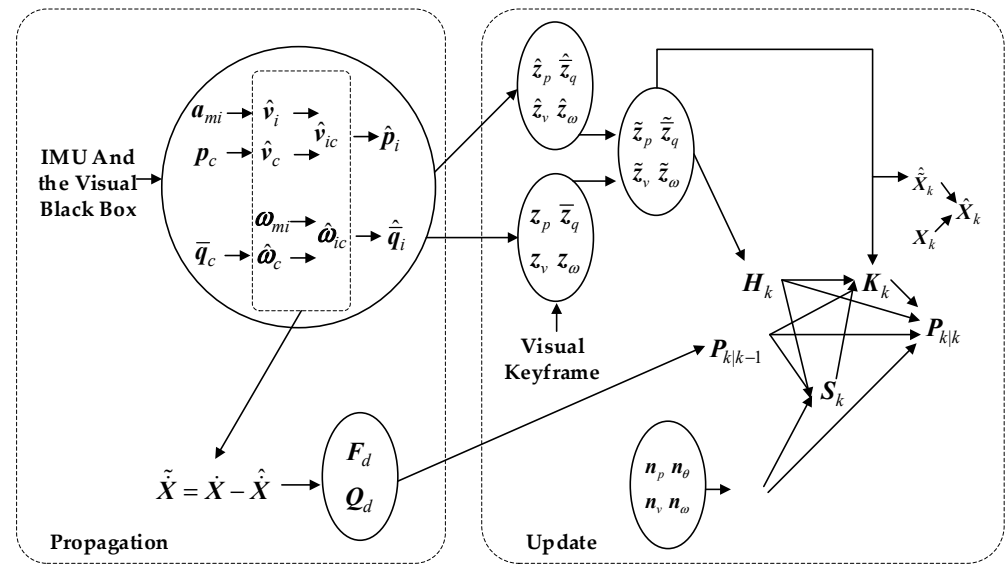


Figure 3. The data flow of the variables in the ICEKF.

4. Nonlinear Observability Analysis

The nonlinear system with an ICEKF can function properly if it has local weak observability, as described in [27]. To simplify the observability analysis with angular velocity layer fusion, a pair of virtual coupled measurement variables are defined as angular velocity ω_{mic} with bias $b_{\omega_{ic}}$, which can be regarded as the yields obtained from correcting ω_{mi} and b_{ω_i} with visual measurement. Thus, referring to the observation system modeling in [16], the nonlinear system representing the fusion measurement results can be expressed as follows:

$$\dot{\hat{X}} = \begin{bmatrix} \dot{\hat{p}}_{ic} \\ \dot{\hat{v}}_c \\ \dot{\hat{v}}_i \\ \dot{\hat{q}}_{ic} \\ \dot{\hat{q}}_i \\ \dot{\hat{\omega}}_c \\ \dot{\hat{\omega}}_i \\ \dot{\hat{b}}_{a_i} \\ \dot{\hat{b}}_{\omega_i} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} \overbrace{\mu_v \hat{v}_i + \lambda(1 - \mu_v) R_c^i \hat{v}_c}^{f_0} \\ 0_{3 \times 1} \\ -R_i b_{a_i} - g \\ \frac{1}{2} \Xi(\dot{\hat{q}}_{ic}) b_{\omega_{ic}} \\ \frac{1}{2} \Xi(\dot{\hat{q}}_i) b_{\omega_i} \\ 0_{3 \times 1} \\ 0_{3 \times 1} \\ 0_{3 \times 1} \\ 0_{3 \times 1} \\ 0 \end{bmatrix} + \begin{bmatrix} 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{4 \times 3} \\ \frac{1}{2} \Xi(\dot{\hat{q}}_i) \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0 \end{bmatrix} \omega_{mi} + \begin{bmatrix} 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ \frac{1}{2} \Xi(\dot{\hat{q}}_{ic}) \\ 0_{4 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0 \end{bmatrix} \omega_{mic} + \begin{bmatrix} 0_{3 \times 3} \\ 0_{3 \times 3} \\ R_i \\ 0_{4 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0_{3 \times 3} \\ 0 \end{bmatrix} a_{mi} \quad (33)$$

where for a general unit quaternion \bar{q} , there is the following:

$$\begin{cases} \dot{\hat{q}} = 0.5 \Xi(\dot{\hat{q}}) \omega \\ \Xi(\bar{q}) = \begin{bmatrix} -q^T \\ q_0 I_3 - [q \times] \end{bmatrix} \end{cases}, \quad (34)$$

The measurement functions are designed as $\mathbf{h}(\dot{\mathbf{X}}) = [\mathbf{h}_1^T, \dots, \mathbf{h}_8^T]^T$, with $\mathbf{h}_1 = (\mathbf{p}_{ic} - \mathbf{R}_{ic}\mathbf{p}_i^c)\lambda$, $\mathbf{h}_2 = \bar{\mathbf{q}}_{ic}$, $\mathbf{h}_3 = \bar{\mathbf{q}}_i$, $\mathbf{h}_4 = \bar{\mathbf{q}}_{ic}^T \bar{\mathbf{q}}_{ic}$, $\mathbf{h}_5 = \bar{\mathbf{q}}_i^T \bar{\mathbf{q}}_i$, $\mathbf{h}_6 = \mathbf{v}_c$, $\mathbf{h}_7 = \boldsymbol{\omega}_c$, and $\mathbf{h}_8 = \boldsymbol{\omega}_i$.

According to the detailed deduction in Appendix C, the observability matrix of the system is constructed with the Lie derivative as follows:

$$\mathbf{\Omega} = \begin{bmatrix} \nabla L^0 \mathbf{h}_1 \\ \nabla L^0 \mathbf{h}_2 \\ \nabla L^0 \mathbf{h}_3 \\ \nabla L^0 \mathbf{h}_4 \\ \nabla L^0 \mathbf{h}_5 \\ \nabla L^0 \mathbf{h}_6 \\ \nabla L^0 \mathbf{h}_7 \\ \nabla L^0 \mathbf{h}_8 \\ \nabla L_{f_0}^1 \mathbf{h}_1 \\ \nabla L_{f_0}^1 \mathbf{h}_3 \\ \nabla L_{f_0}^2 \mathbf{h}_1 \\ \nabla L_{f_3}^1 L_{f_0}^1 \mathbf{h}_1 \end{bmatrix} = \begin{bmatrix} \underbrace{\mathbf{U}_1}_{p_{ic}} & \underbrace{\mathbf{0}}_{v_c} & \underbrace{\mathbf{0}}_{v_i} & \underbrace{\mathbf{U}_2}_{\bar{\mathbf{q}}_{ic}} & \underbrace{\mathbf{0}}_{\bar{\mathbf{q}}_i} & \underbrace{\mathbf{0}}_{\omega_c} & \underbrace{\mathbf{0}}_{\omega_i} & \underbrace{\mathbf{0}}_{b_a} & \underbrace{\mathbf{0}}_{b_\omega} & \underbrace{\mathbf{U}_3}_{\lambda} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_4 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_5 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_6 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_6 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_8 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_9 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_{10} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_{11} & \mathbf{U}_{12} & \mathbf{G}_{[9,4]} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_{[9,8]} & \mathbf{0} & \mathbf{G}_{[9,10]} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_{13} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_{[11,4]} & \mathbf{G}_{[11,5]} & \mathbf{0} & \mathbf{0} & \mathbf{U}_{15} & \mathbf{G}_{[11,9]} & \mathbf{G}_{[11,10]} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G}_{[12,4]} & \mathbf{G}_{[12,5]} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_{16} \end{bmatrix}_{38 \times 30}, \quad (35)$$

where each column corresponds to the entries in the state vector of the ICEKF, the matrices \mathbf{G} with indices represent blocks that are irrelevant to the rank analysis, and the matrices \mathbf{U} with subscripts are blocks contributing to the column rank of $\mathbf{\Omega}$, as follows: $\mathbf{U}_1 = \mathbf{I}_3\lambda$, $\mathbf{U}_4, \mathbf{U}_5 = \mathbf{I}_4$, $\mathbf{U}_8, \mathbf{U}_9, \mathbf{U}_{10} = \mathbf{I}_3$, $\mathbf{U}_{12} = \lambda\mu_v\mathbf{I}_3$, $\mathbf{U}_{14} = 0.5\Xi(\dot{\bar{\mathbf{q}}}_i)$, $\mathbf{U}_{15} = -\lambda\mu_v\mathbf{R}_i$, and $\mathbf{U}_{16} = \mathbf{U}(\mathbf{R}_i)$.

To prove that $\mathbf{\Omega}$ has full rank, block Gaussian elimination is applied. The rows of blocks can be rearranged so that only one block in each row is allowed to determine whether the corresponding column of blocks has full column rank. Following this process, all the block columns in $\mathbf{\Omega}$, with the exception of the last one corresponding to \mathbf{U}_{16} , have full column rank when λ, μ_v are nonzero.

Expanding the rotation matrix \mathbf{R}_i as three columns yields the following:

$$L_{f_3}^1 L_{f_0}^1 \mathbf{h}_1 = \lambda\mu_v \mathbf{R}_i = \lambda\mu_v [\mathbf{r}_x \mathbf{r}_y \mathbf{r}_z], \quad (36)$$

An explicit form is achieved by inspecting the Lie derivative of Equation (36) as $\mathbf{U}_{16} = \mu_v [\mathbf{r}_x^T \mathbf{r}_y^T \mathbf{r}_z^T]^T_{9 \times 1}$. As described in [16], if the linear acceleration was excited on at least one axis, one of $\mathbf{r}_x, \mathbf{r}_y$, or \mathbf{r}_z would be a nonzero vector.

In accordance with the analysis above, it can be demonstrated that for the visual-inertial system with the ICEKF described in this paper, when λ, μ_v are nonzero and the IMU is excited in any direction, the observability matrix has full column rank, which means that the system has local weak observability [27]. The above conditions are readily fulfilled during practical applications.

5. Simulation and Experiments

Three-dimensional motion simulation and dataset experiments are conducted in this section to analyze the performance of the ICEKF.

5.1. Simulation

Assuming that the coordinate frames of the IMU and the camera are aligned, the pose transformation between them can be omitted. To observe the convergence process, we

deliberately assign random initial values to the prior covariance matrix $P_{k|k}$. The starting point is [1,0,0]. The covariance matrix of measurement error R_k is designed as a skew-symmetric matrix with small random values. Gaussian noise is introduced to all the virtual measurement values. The coefficients are set as $\mu_v = 0.5$, $\mu_\omega = 0.5$, and $\lambda = 1$. The virtual IMU-camera rig is directed to move along a virtual helical trajectory, as described by the following equation:

$$\begin{cases} x = \cos\left(\frac{1}{2}\pi t\right) \\ y = \sin\left(\frac{1}{2}\pi t\right) \\ z = t \end{cases}, \tag{37}$$

The three-dimensional position curves of the ICEKF and the ground truth are shown in Figure 4, which reveals that the estimated results align closely with the true values. The position errors of the ICEKF estimation against the ground truth of the IMU-camera rig along the three axes are depicted in Figure 5, and Figure 6 shows the orientation errors. The curve of λ is shown in Figure 7.

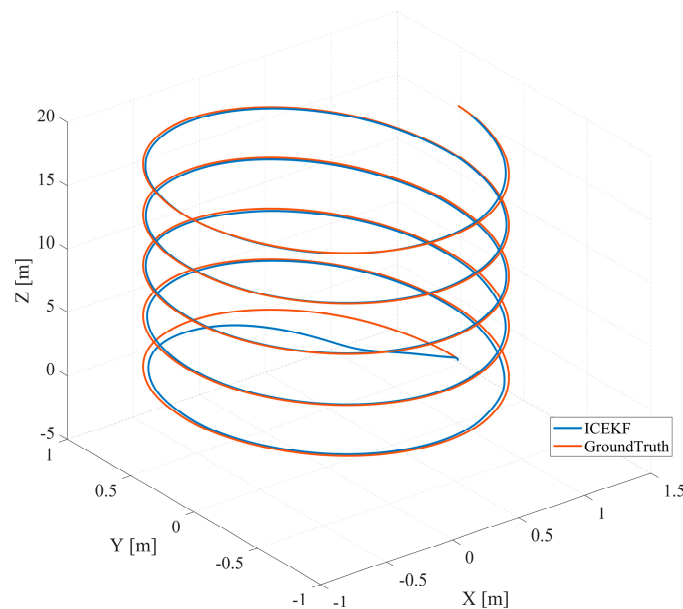


Figure 4. Three-dimensional position curves of the simulation.

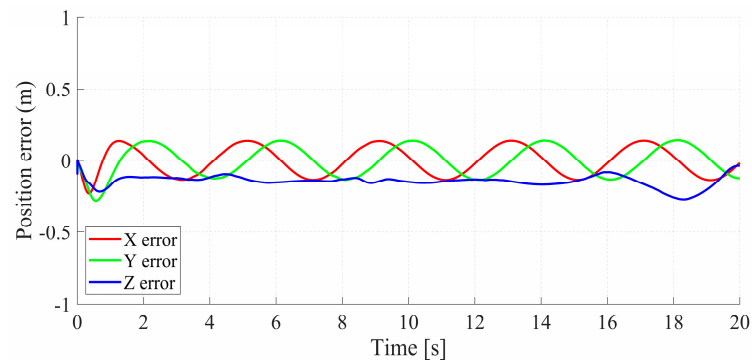


Figure 5. Position error of the simulation.

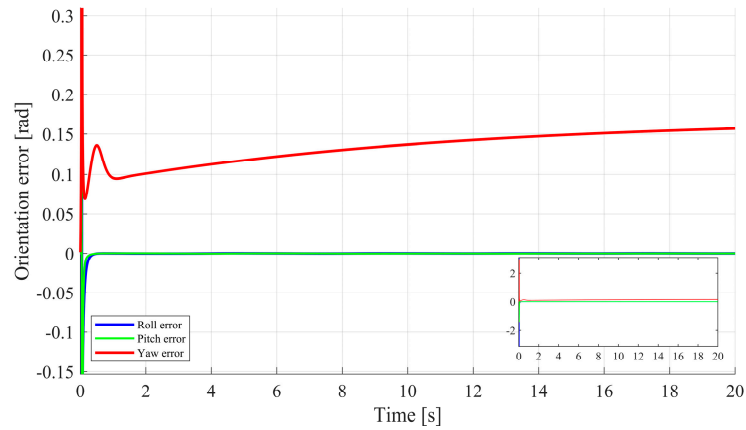


Figure 6. Orientation error of the simulation.

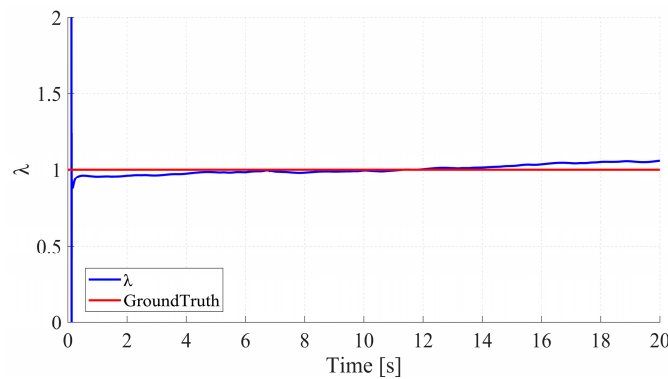


Figure 7. Visual scale of the simulation.

The norm errors between the ground truth and the simulation results, including the statistical values of the RMSE (root mean square error), mean error, and STD (standard deviation) for both coupled translation p_{ic} and attitude \bar{q}_{ic} across all three axes, are detailed in Table 2. The errors of the attitude are obtained by converting \bar{q}_{ic} to roll-pitch-yaw.

Table 2. Norm errors of the translation and attitude between the ground truth and the simulation results.

Translation RMSE	Translation Mean Error	Translation STD	Attitude RMSE	Attitude Mean Error	Attitude STD
0.207 m	0.1447 m	0.1408 m	0.1684 rad	0.1348 rad	0.1008 rad

The curves and numerical statistics show that despite the randomized $P_{k|k}$ and R_k alongside a high yaw angle velocity disturbing the initial convergence, the ICEKF ultimately converges without much specific parameter tuning. This indicates that the ICEKF possesses strong robustness while yielding reliable motion estimates.

To further test the filter’s ability to converge, an initial state estimate set is designed which conducts the virtual rig starting from the different three-dimensional points to the actual desired starting point [1,0,0]. The initial estimate set, varying from [0,−1,−1] to [1,0,0], and the converging process are illustrated by Figure 8, which depicts how even when facing uncertain initial estimates, the filter can converge with a decent performance.

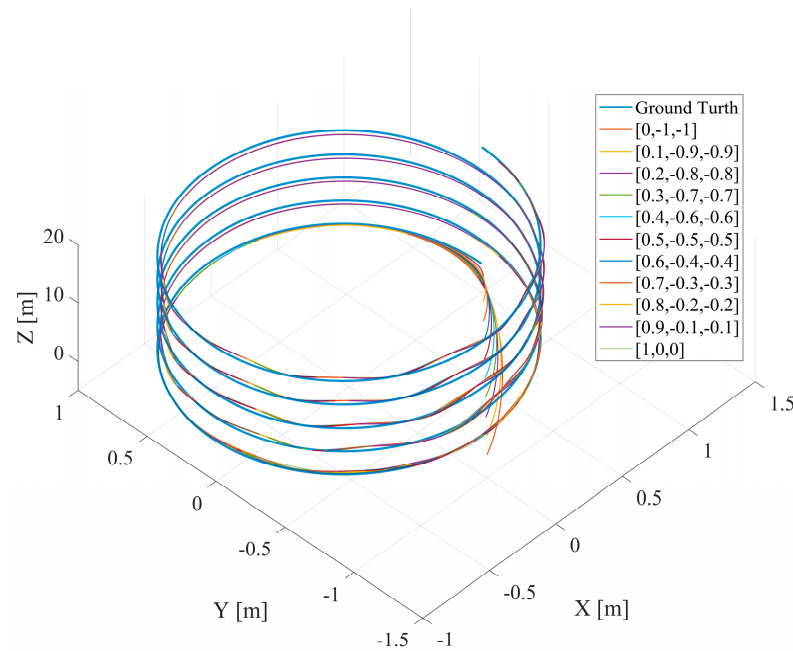


Figure 8. The initial estimate set and the converging process.

5.2. Dataset Experiment

To inspect the effectiveness of the algorithm in real-world scenarios, this paper employs the EuRoC ROOM01 and MainHall02 datasets [28] for validation. Since this study investigates the performance of the loosely coupled visual–inertial algorithm itself instead of visual odometry, one of the state-of-the-art visual SLAM algorithms, monocular ORB-SLAM V3 [29], serves as the visual black-box approach. Given that this algorithm has been confirmed to exhibit exceptional ATE performance, this paper focuses on RPE comparisons of different algorithms by employing EVO tools [30] to analyze trajectories. The ICEKF operates with the IMU data at 200 Hz, and meanwhile the visual measurement runs at 20 Hz. To facilitate the comparison, each frame from the visual measurement with ORB-SLAM is considered a keyframe.

The coefficients are $\mu_v = 0.9$ and $\mu_\omega = 0.5$, and the initial scale factor is determined based on the prior results of the visual algorithm as $\lambda = 2.04$. Given that the hardware setup and scenario are fixed in the dataset, the initial covariance matrix $\mathbf{P}_{k|k}$ of the ICEKF can be easily adjusted via multiple runs. In the first run, the initial $\mathbf{P}_{k|k}$ is selected randomly, and eventually stabilizes as the algorithm progresses. The final updated $\mathbf{P}_{k|k}$ is utilized to run the scenario test again, which greatly increases the convergence speed and provides high resilience against noise. For applications in unknown environments, initial covariance estimation can be conducted in a similar scenario to further improve convergence efficiency.

Dataset experiments with monocular ORB-SLAM, its IMU variant, and the monocular MSCKF with an IMU were performed. The entire experiment processes are shown in Figures 9 and 10.

The contrast between the three-dimensional position output of the ICEKF and the ground truth is shown in Figures 11 and 12. The three-dimensional position errors and the orientation errors between the output of the ICEKF and the ground truth after applying alignment using Umeyama’s method [30] from the two dataset tests are shown in Figures 13–16.

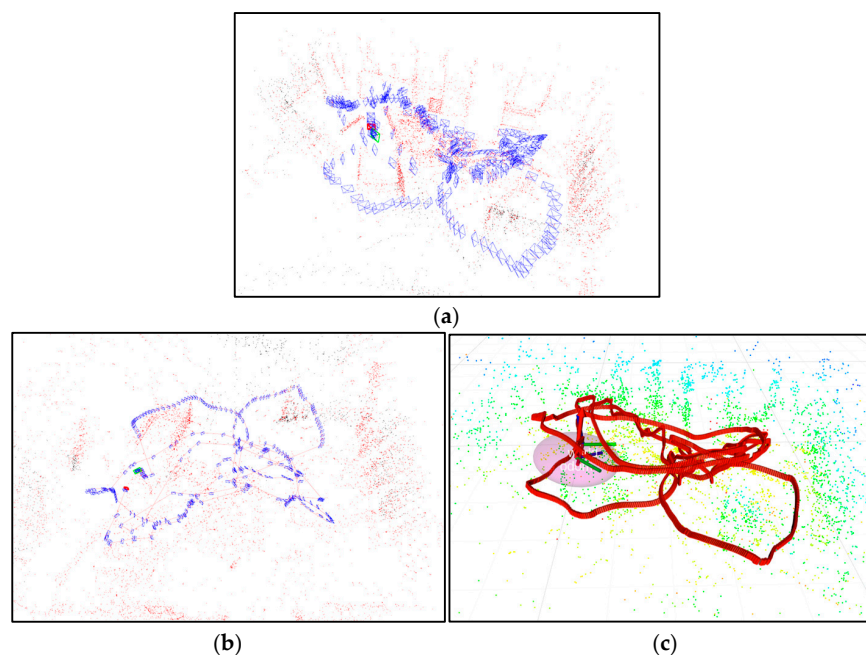


Figure 9. Dataset experiments (ROOM01): (a) monocular ORB-SLAM, (b) monocular ORB-SLAM with IMU, and (c) monocular MSCKF.

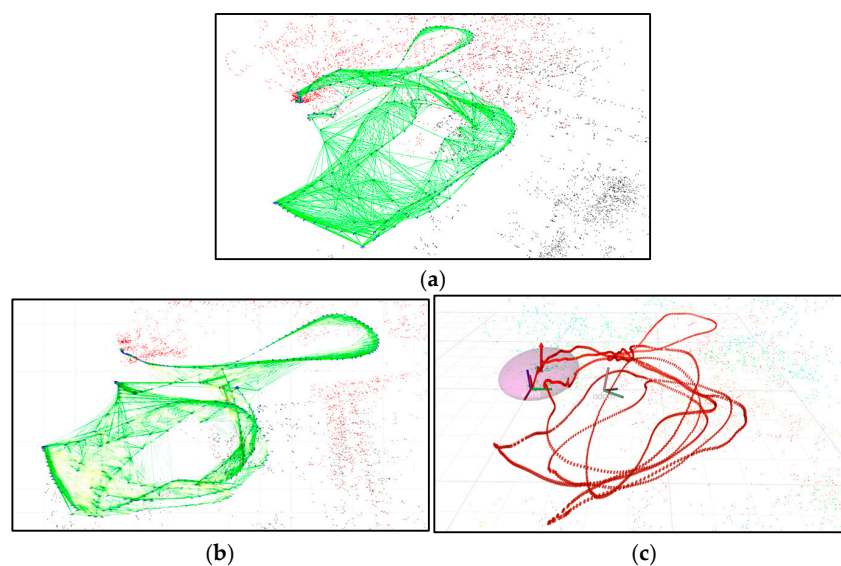


Figure 10. Dataset experiments (MainHall02): (a) monocular ORB-SLAM, (b) monocular ORB-SLAM with IMU, and (c) monocular MSCKF.

Trajectory analysis using EVO is conducted, in which the ICEKF with the results from ORB-SLAM and its IMU variant, as well as the data obtained from the monocular MSCKF [31], are compared against the ground truth, with the statistical RPE results presented in Tables 3 and 4.

The comparative results indicate that the ICEKF ensures real-time measurement through high-frequency IMU updates and achieves an accuracy comparable to that of the monocular ORB-SLAM and its IMU variant while outperforming the monocular MSCKF. Although the trajectory does not reach the high accuracy of ORB-SLAM-based methods owing to the lack of posterior batch optimization processing, the computational complexity of the coupled part of the algorithm is only $O(n)$, which benefits localization applications. Moreover, the algorithm framework can be directly utilized with other localization algorithms to improve their measurement stability in complex environments.

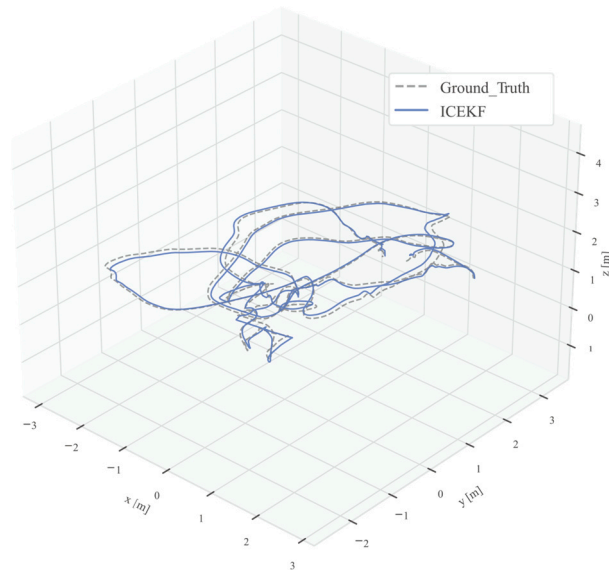


Figure 11. Position output of the ICEKF and the ground truth (ROOM01).

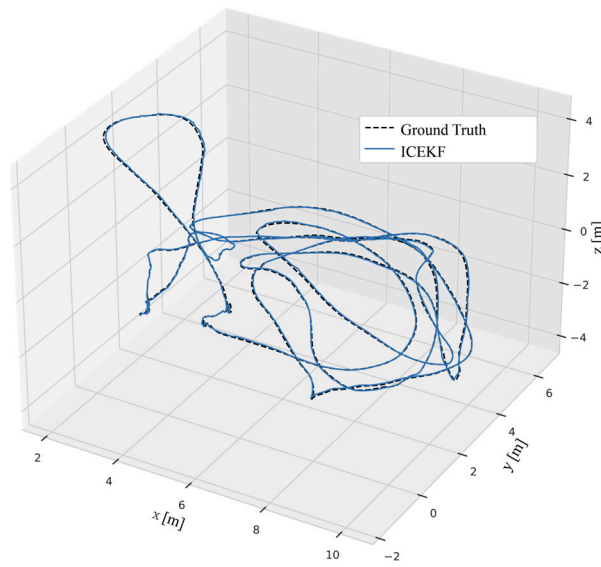


Figure 12. Position output of the ICEKF and the ground truth (MainHall02).

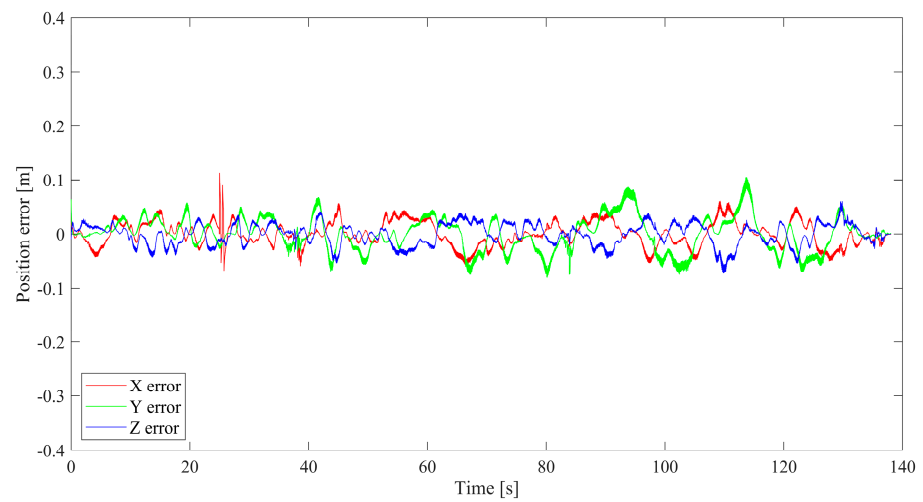


Figure 13. Position error between the ICEKF and the ground truth (ROOM01).

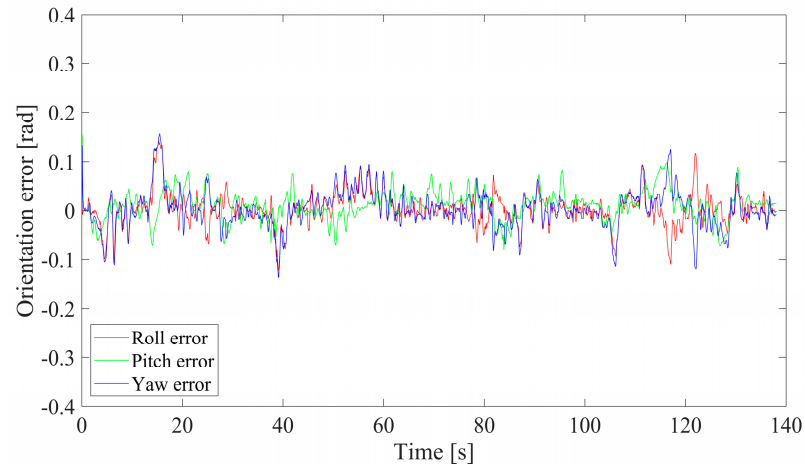


Figure 14. Orientation error between the ICEKF and the ground truth (ROOM01).

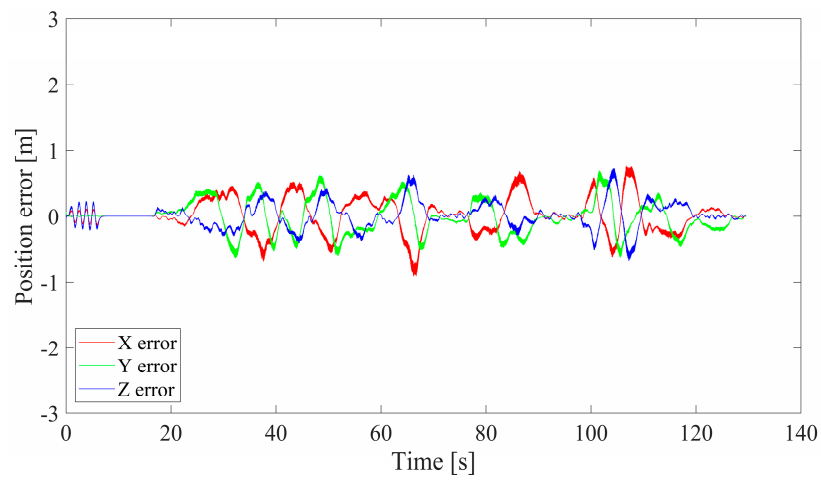


Figure 15. Position error between the ICEKF and the ground truth (MainHall02).

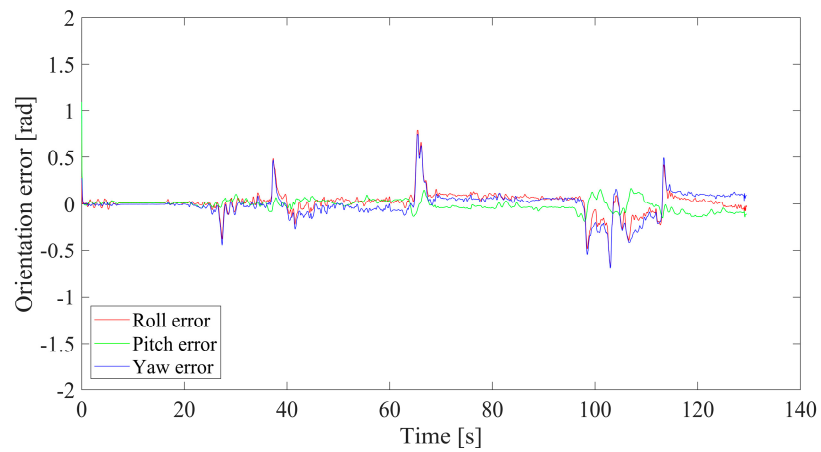


Figure 16. Orientation error between the ICEKF and the ground truth (MainHall02).

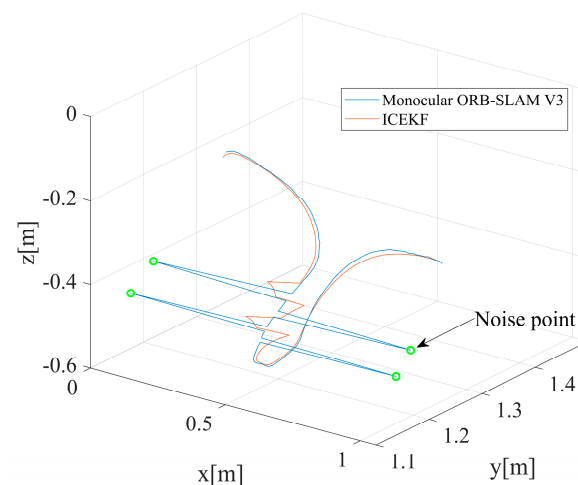
Table 3. Norm errors of the translations between the ground truth and the results of the dataset experiments (ROOM01).

	Translation RMSE	Translation Mean Error	Translation STD
ICEKF	0.04153 m	0.00574 m	0.04112 m
Monocular ORB-SLAM V3	0.0393 m	0.0355 m	0.01678 m
Monocular ORB-SLAM V3 with IMU	0.003841 m	0.002973 m	0.002433 m
Monocular MSCKF	0.1305 m	0.05366 m	0.119 m

Table 4. Norm errors of the translations between the ground truth and the results of the dataset experiments (MainHall02).

	Translation RMSE	Translation Mean Error	Translation STD
ICEKF	0.08922 m	0.009765 m	0.08869 m
Monocular ORB-SLAM V3	0.735 m	0.533 m	0.506 m
Monocular ORB-SLAM V3 with IMU	0.001366 m	0.004981 m	0.01272 m
Monocular MSCKF	0.2689 m	0.09905 m	0.25 m

To assess the robustness of the ICEKF with regard to leap noise from visual measurements, this study incorporates random stimulations with an amplitude of 0.5 m into the visual measurement result to imitate harsh visual measurement failure. Figure 17 shows a comparison of the partial trajectory from the ICEKF against the corrupted visual measurement. This indicates that the ICEKF can significantly mitigate the impact of sporadic instability from visual measurements.

**Figure 17.** Comparison of the partial trajectory from the ICEKF against the visual measurement with leap noise.

6. Conclusions

This study introduces a novel weighted loosely coupled algorithm for fusing data from IMUs and visual measurements and demonstrates its observability. The algorithm, of which the coupled part possesses a computational complexity of $O(n)$, retains the high-

frequency update capability inherent to the EKF framework, thus enabling high-accuracy active localization. Both simulation and dataset tests reveal that the ICEKF achieves deep integration of the IMU and the visual data on the velocity layer throughout the updating process, which benefits the localization performance. By conveniently tuning the weights assigned to different data sources, the framework improves both the fusion accuracy and the resilience to abrupt noise. In future studies, we will aim to apply the approach across multiple devices by analyzing the sensitivity of the coupling coefficients in challenging real-life scenarios.

Author Contributions: Conceptualization, Z.L. and P.T.; methodology, C.L.; software, Z.L. and T.W.; validation, C.L.; formal analysis, C.L.; investigation, Z.L. and P.T.; resources, Z.L. and P.T.; data curation, C.L. and T.W.; writing—original draft preparation, Z.L. and C.L.; writing—review and editing, T.W. and P.T.; visualization, Z.L. and C.L.; supervision, P.T.; project administration, Z.L. and P.T.; funding acquisition, T.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was sponsored by the National Key Research and Development Program of China [Grant number 2022YFC3320503] and the Foundation for Innovative Research Groups of the National Natural Science Foundation of China [Grant number 12221002], funded by Tao Wang.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The source code presented in this study is available upon request from the corresponding author.

Acknowledgments: The authors acknowledge the anonymous reviewers for their helpful comments on the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Assuming that the rotation \bar{q} is accomplished during a certain period and that the corresponding body angular velocity is ω , $\dot{\bar{q}} = \frac{1}{2}\bar{q} \otimes \bar{\omega}$ [30], where $\bar{\omega} = [0, \omega]$. The μ -th power of \bar{q} is written as \bar{q}^μ , which is a unit quaternion that denotes scaling the rotation angle around the virtual axis with $\mu \in [0, 1]$ [32].

A three-element vector p rotating according to \bar{q} is written as $\bar{q} \otimes \bar{p} \otimes \bar{q}^* \triangleq R(\bar{q})p$ with $\bar{p} = [0, p^T]^T$.

For general quaternions, the error between the measurement and the expectation is defined as $\delta\bar{q}$ from $\bar{q} = \hat{q} \otimes \delta\bar{q}$. The derivative form is $\delta\dot{\bar{q}} = \dot{\hat{q}}^* \otimes \left(\dot{\bar{q}} - \dot{\hat{q}} \otimes \delta\bar{q} \right)$ [16].

According to $R(\delta\bar{q}) \approx I_3 + [\delta\theta \times]$ [23], there is $R(\bar{q}) = \hat{R}(\hat{q})R(\delta\bar{q}) \approx \hat{R}(\hat{q})(I_3 + [\delta\theta \times])$.

Assuming that there are two quaternions, $\bar{q} = [q_0, q^T]^T$ and $\bar{p} = [p_0, p^T]^T$, the following equation exists:

$$\bar{q} \otimes \bar{p} = \begin{bmatrix} q_0 & -q^T \\ q & q_0 I_3 + [q \times] \end{bmatrix} \otimes \bar{p} = \begin{bmatrix} p_0 & -p^T \\ p & p_0 I_3 + [p \times] \end{bmatrix} \otimes \bar{q}. \quad (A1)$$

Appendix B

The following lemmas are proposed in this paper for the coupling process on the velocity layer.

Lemma A1. For the IMU-camera measurement system in Figure 1, let $\bar{q}_w^{c'}$ be the equivalent rotation of \bar{q}_w^c described in the IMU frame. Assuming that \bar{q}_w^{ic} , \bar{q}_w^i and $\bar{q}_w^{c'}$ are unit quaternions with $\bar{q}_w^{ic} = \bar{q}_w^{c'} = \bar{q}_w^i$, $\bar{q}_w^{ic} = (\bar{q}_w^i)^{\mu_\omega} \otimes (\bar{q}_w^{c'})^{1-\mu_\omega}$, where $\mu_\omega \in [0, 1]$.

Proof of Lemma A1. Rewriting the quaternions into axis-angle form [32] yields $\bar{q}_w^{ic} = [\cos \theta_{ic}, \mathbf{r}_{ic} \sin \theta_{ic}]$, $\bar{q}_w^{c'} = [\cos \theta_c, \mathbf{r}_c \sin \theta_c]$, and $\bar{q}_w^i = [\cos \theta_i, \mathbf{r}_i \sin \theta_i]$, with $\theta_c = \theta_i = \theta_{ic} = \theta$, and unit vectors $\mathbf{r}_{ic} = \mathbf{r}_c = \mathbf{r}_i = \mathbf{r}$.

Inspecting the exponential form of quaternions [33], we have the following:

$$\begin{aligned} (\bar{q}_w^{c'})^{(1-\mu_\omega)} &= \exp((1-\mu_\omega) \log \bar{q}_w^{c'}) = \exp((1-\mu_\omega)[0, \theta \mathbf{r}]) \quad (A2) \\ &= [\cos((1-\mu_\omega)\theta), \mathbf{r} \sin((1-\mu_\omega)\theta)] \end{aligned}$$

$$\begin{aligned} (\bar{q}_w^i)^{\mu_\omega} &= \exp(\mu_\omega \log \bar{q}_w^i) = \exp(\mu_\omega[0, \theta \mathbf{r}]) \quad (A3) \\ &= [\cos(\mu_\omega\theta), \mathbf{r} \sin(\mu_\omega\theta)] \end{aligned}$$

Three conditions are discussed as follows:

1. When $\mu_\omega = 1$, $(\bar{q}_w^{c'})^0 = [1, 0]$, and $(\bar{q}_w^i)^1 = [\cos \theta, \mathbf{r} \sin \theta]$, $(\bar{q}_w^{c'})^0 \otimes (\bar{q}_w^i)^1 = [\cos \theta, \mathbf{r} \sin \theta] = \bar{q}_w^{ic}$.
2. When $\mu_\omega = 0$, similarly, $(\bar{q}_w^{c'})^1 \otimes (\bar{q}_w^i)^0 = \bar{q}_w^{ic}$.
3. When $\mu_\omega \in (0, 1)$, there is the following:

$$\begin{aligned} (\bar{q}_w^i)^{\mu_\omega} \otimes (\bar{q}_w^{c'})^{(1-\mu_\omega)} &= [\cos((1-\mu_\omega)\theta), \mathbf{r} \sin((1-\mu_\omega)\theta)][\cos(\mu_\omega\theta), \mathbf{r} \sin(\mu_\omega\theta)] \\ &= [\cos((1-\mu_\omega)\theta) \cos(\mu_\omega\theta) - \sin((1-\mu_\omega)\theta) \sin(\mu_\omega\theta) (\mathbf{r} \cdot \mathbf{r}), \\ &\quad \mathbf{r} \cos((1-\mu_\omega)\theta) \sin(\mu_\omega\theta) + \mathbf{r} \cos(\mu_\omega\theta) \sin((1-\mu_\omega)\theta) + (\mathbf{r} \times \mathbf{r}) \sin((1-\mu_\omega)\theta) \sin(\mu_\omega\theta)] \quad (A4) \\ &= [\cos \theta, \mathbf{r} \sin \theta] = \bar{q}_w^{ic} \end{aligned}$$

Therefore, for all three conditions, $\bar{q}_w^{ic} = (\bar{q}_w^i)^{\mu_\omega} \otimes (\bar{q}_w^{c'})^{1-\mu_\omega}$ holds. \square

Lemma A2. Letting $\mu \in [0, 1]$, for the derivative of the μ -th power of a general unit quaternion \bar{q} , $\frac{d}{dt}(\bar{q}^\mu) = \frac{1}{2}\bar{q}^\mu \otimes (\mu\bar{\omega})$.

Proof of Lemma A2. Referring to Equation (314) in [32] (pp. 483), for a general quaternion in the space-reference inertia axes, namely, the body frame in this paper, $\frac{d}{dt}(\bar{q}) = \frac{1}{2}\bar{q} \otimes (\bar{\omega})$.

Referring to Equation (15) in [34] (pp. 168), for a unit quaternion \bar{q} denoting rotation α around unit vector \mathbf{r} , $\frac{d}{dt}\alpha = \boldsymbol{\omega} \cdot \mathbf{r}$ exists.

To scale the rotation angle with a coefficient $\mu \in [0, 1]$, since the vector \mathbf{r} remains unchanged, $\frac{d}{dt}\mu\alpha = (\mu\boldsymbol{\omega}) \cdot \mathbf{r}$. Thus, $\frac{d}{dt}(\bar{q}^\mu) = \frac{1}{2}\bar{q}^\mu \otimes (\mu\bar{\omega})$ holds. \square

Appendix C

The Lie derivative of the measurement function $\mathbf{h}(x)$ with respect to the vector $\mathbf{f}(x)$ is written as follows:

$$L_f \mathbf{h}(x) = \nabla_f \mathbf{h}(x) = \frac{\partial \mathbf{h}(x)}{\partial x} \mathbf{f}(x). \quad (A5)$$

The k -th order Lie derivative of $\mathbf{h}(x)$ with respect to $\mathbf{f}(x)$ is written as follows:

$$L_f^k \mathbf{h}(x) = \frac{\partial (L_f^{k-1} \mathbf{h}(x))}{\partial x} \mathbf{f}(x). \quad (A6)$$

Specifically, the zeroth-order Lie derivative of $\mathbf{h}(\mathbf{x})$ is the measurement function itself, namely $L^0\mathbf{h}(\mathbf{x}) \triangleq \mathbf{h}(\mathbf{x})$.

By deducting the gradient of the zeroth-order Lie derivatives of the measurement function among $\mathbf{h}(\dot{\mathbf{X}}) = [\mathbf{h}_1^T, \dots, \mathbf{h}_8^T]^T$, we obtain the following:

$$\left\{ \begin{array}{l} \nabla L^0\mathbf{h}_1 = [\mathbf{I}_3\lambda \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \lambda\Gamma(\bar{\mathbf{q}}_{ic}, \mathbf{p}_i^c) \quad \mathbf{0}_{3\times4} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3}\mathbf{0}_{3\times3} \quad (\mathbf{p}_{ic} - \mathbf{R}_{ic}\mathbf{p}_i^c)] \\ \nabla L^0\mathbf{h}_2 = [\mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{I}_4 \quad \mathbf{0}_{4\times4} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times1}] \\ \nabla L^0\mathbf{h}_3 = [\mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times4} \quad \mathbf{I}_4 \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times3} \quad \mathbf{0}_{4\times1}] \\ \nabla L^0\mathbf{h}_4 = \begin{bmatrix} \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & 2(\bar{\mathbf{q}}_{ic})^T & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0} \end{bmatrix} \\ \nabla L^0\mathbf{h}_5 = \begin{bmatrix} \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times4} & 2(\bar{\mathbf{q}}_i)^T & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0}_{1\times3} & \mathbf{0} \end{bmatrix} \\ \nabla L^0\mathbf{h}_6 = [\mathbf{0}_{3\times3} \quad \mathbf{I}_3 \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times4} \quad \mathbf{0}_{3\times4} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times1}] \\ \nabla L^0\mathbf{h}_7 = [\mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times4} \quad \mathbf{0}_{3\times4} \quad \mathbf{I}_3 \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times1}] \\ \nabla L^0\mathbf{h}_8 = [\mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times4} \quad \mathbf{0}_{3\times4} \quad \mathbf{0}_{3\times3} \quad \mathbf{I}_3 \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times3} \quad \mathbf{0}_{3\times1}] \end{array} \right. \quad (A7)$$

The first-order Lie derivative of \mathbf{h}_1 with respect to \mathbf{f}_0 and its gradient are as follows:

$$L_{\mathbf{f}_0}^1\mathbf{h}_1 = \nabla L^0\mathbf{h}_1 \cdot \mathbf{f}_0 = \lambda(\mu_v\mathbf{v}_i + \lambda(1 - \mu_v)\mathbf{R}_c^i\mathbf{v}_c) + \frac{1}{2}\lambda\Gamma(\bar{\mathbf{q}}_{ic}, \mathbf{p}_i^c)\Xi(\dot{\bar{\mathbf{q}}}_{ic})\mathbf{b}_{wi}. \quad (A8)$$

$$\nabla L_{\mathbf{f}_0}^1\mathbf{h}_1 = \begin{bmatrix} \mathbf{0}_{3\times3} & \lambda^2(1 - \mu_v)\mathbf{R}_c^i & \lambda\mu_v\mathbf{I}_3 & \mathbf{G}_{[9,4]} & \mathbf{0}_{3\times4} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{G}_{[9,8]} & \mathbf{0}_{3\times3} & \mathbf{G}_{[9,10]} \end{bmatrix} \quad (A9)$$

The first-order Lie derivative of \mathbf{h}_3 with respect to \mathbf{f}_0 and its gradient are as follows:

$$L_{\mathbf{f}_0}^1\mathbf{h}_3 = \nabla L^0\mathbf{h}_3 \cdot \mathbf{f}_0 = 0.5\Xi(\dot{\bar{\mathbf{q}}}_i)\mathbf{b}_{wi}. \quad (A10)$$

$$\nabla L_{\mathbf{f}_0}^1\mathbf{h}_3 = \begin{bmatrix} \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times4} & \mathbf{0}_{4\times4} & \mathbf{I}_3 & \mathbf{0}_{4\times3} & \mathbf{0}_{4\times3} & 0.5\Xi(\dot{\bar{\mathbf{q}}}_i) & \mathbf{0}_{4\times1} \end{bmatrix} \quad (A11)$$

The second-order Lie derivative of \mathbf{h}_1 with respect to \mathbf{f}_0 and its gradient are as follows:

$$L_{\mathbf{f}_0}^2\mathbf{h}_1 = \nabla L_{\mathbf{f}_0}^1\mathbf{h}_1 \cdot \mathbf{f}_0 = \lambda\mu_v(-\mathbf{R}_i\mathbf{b}_{ai} - \mathbf{g}) + \frac{1}{2}\mathbf{G}_{[9,4]}\Xi(\dot{\bar{\mathbf{q}}}_{ic})\mathbf{b}_{wi}. \quad (A12)$$

$$\nabla L_{\mathbf{f}_0}^2\mathbf{h}_1 = \begin{bmatrix} \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{G}_{[11,4]} & \mathbf{G}_{[11,5]} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & -\lambda\mu_v\mathbf{R}_i & \mathbf{G}_{[11,9]} & \mathbf{G}_{[11,10]} \end{bmatrix} \quad (A13)$$

The second-order Lie derivative of \mathbf{h}_1 with respect to \mathbf{f}_0 as well as \mathbf{f}_3 and its gradient are as follows:

$$L_{\mathbf{f}_3}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 = \nabla L_{\mathbf{f}_0}^1\mathbf{h}_1 \cdot \mathbf{f}_3 = \lambda\mu_v\mathbf{R}_i = \left[L_{\mathbf{f}_{3,1}}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 \quad L_{\mathbf{f}_{3,2}}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 \quad L_{\mathbf{f}_{3,3}}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 \right]_{3\times3}. \quad (A14)$$

$$\begin{aligned} \nabla L_{\mathbf{f}_3}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 &= \begin{bmatrix} \nabla L_{\mathbf{f}_{3,1}}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 \\ \nabla L_{\mathbf{f}_{3,2}}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 \\ \nabla L_{\mathbf{f}_{3,3}}^1 L_{\mathbf{f}_0}^1\mathbf{h}_1 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{G}_{[12,4]} & \mathbf{G}_{[12,5]} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{0}_{9\times3} & \mathbf{U}(\mathbf{R}_i) \end{bmatrix}_{9\times30} \end{aligned} \quad (A15)$$

References

1. Servières, M.; Renaudin, V.; Dupuis, A.; Antigny, N. Visual and Visual-Inertial SLAM: State of the Art, Classification, and Experimental Benchmarking. *J. Sens.* **2021**, *2021*, 2054828. [CrossRef]
2. He, M.; Zhu, C.; Huang, Q.; Ren, B.; Liu, J. A review of monocular visual odometry. *Vis. Comput.* **2020**, *36*, 1053–1065. [CrossRef]
3. Nangung, H.; Kim, J.S. Collision risk inference system for maritime autonomous surface ships using COLREGs rules compliant collision avoidance. *IEEE Access* **2019**, *9*, 7823–7835. [CrossRef]
4. Lin, Y.; Gao, F.; Qin, T.; Gao, W.; Liu, T.; Wu, W.; Yang, Z.; Shen, S. Autonomous aerial navigation using monocular visual-inertial fusion. *J. Field Robot.* **2018**, *35*, 23–51. [CrossRef]

5. Nemeč, D.; Šimák, V.; Janota, A.; Hruboš, M.; Bubeníková, E. Precise localization of the mobile wheeled robot using sensor fusion of odometry, visual artificial landmarks and inertial sensors. *Robot. Auton. Syst.* **2019**, *112*, 168–177. [[CrossRef](#)]
6. Li, Z.; You, B.; Ding, L.; Gao, H.; Huang, F. Trajectory Tracking Control for WMRs with the Time-Varying Longitudinal Slippage Based on a New Adaptive SMC Method. *Int. J. Aerosp. Eng.* **2019**, *2019*, 4951538. [[CrossRef](#)]
7. Namgung, H. Local route planning for collision avoidance of maritime autonomous surface ships in compliance with COLREGs rules. *Sustainability* **2021**, *14*, 198. [[CrossRef](#)]
8. Alatisé, M.B.; Hancke, G.P. A review on challenges of autonomous mobile robot and sensor fusion methods. *IEEE Access* **2020**, *8*, 39830–39846. [[CrossRef](#)]
9. Tonini, A.; Castelli, M.; Bates, J.S.; Lin, N.N.N.; Painho, M. Visual-Inertial Method for Localizing Aerial Vehicles in GNSS-Denied Environments. *Appl. Sci.* **2024**, *14*, 9493. [[CrossRef](#)]
10. Hou, Z.; Wang, R. A Loosely-Coupled GNSS-Visual-Inertial Fusion for State Estimation Based on Optimization. In Proceedings of the 2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC), Greenville, SC, USA, 12–14 November 2021; pp. 163–168.
11. Talebi, S.P.; Mandic, D.P. On the Dynamics of Multiagent Nonlinear Filtering and Learning. In Proceedings of the 2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP), London, UK, 22–25 September 2024; pp. 1–6.
12. He, X.; Li, B.; Qiu, S.; Liu, K. Visual-Inertial Odometry of Structured and Unstructured Lines Based on Vanishing Points in Indoor Environments. *Appl. Sci.* **2024**, *14*, 1990. [[CrossRef](#)]
13. Sun, Z.; Gao, W.; Tao, X.; Pan, S.; Wu, P.; Huang, H. Semi-Tightly Coupled Robust Model for GNSS/UWB/INS Integrated Positioning in Challenging Environments. *Remote Sens.* **2024**, *16*, 2108. [[CrossRef](#)]
14. Gopaul, N.S.; Wang, J.; Hu, B. Loosely coupled visual odometry aided inertial navigation system using discrete extended Kalman filter with pairwise time correlated measurements. In Proceedings of the 2017 Forum on Cooperative Positioning and Service (CPGPS), Harbin, China, 19–21 May 2017; pp. 283–288.
15. Weiss, S.M. Vision Based Navigation for Micro Helicopters. Ph.D. Thesis, ETH Zurich, Zurich, Switzerland, 2012.
16. Kelly, J.; Sukhatme, G.S. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. *Int. J. Robot. Res.* **2011**, *30*, 56–79. [[CrossRef](#)]
17. Weiss, S.; Siegwart, R. Real-time metric state estimation for modular vision-inertial systems. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 4531–4537.
18. Achtelik, M.W.; Weiss, S.; Chli, M.; Dellaert, F.; Siegwart, R. Collaborative stereo. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 2242–2248.
19. Brossard, M.; Bonnabel, S.; Barrau, A. Invariant Kalman filtering for visual inertial SLAM. In Proceedings of the 2018 21st International Conference on Information Fusion (FUSION), Cambridge, UK, 10–13 July 2018; pp. 2021–2028.
20. Sun, W.; Li, Y.; Ding, W.; Zhao, J. A Novel Visual Inertial Odometry Based on Interactive Multiple Model and Multi-state Constrained Kalman Filter. *IEEE Trans. Instrum. Meas.* **2023**, *73*, 5000110. [[CrossRef](#)]
21. Fornasier, A.; Ng, Y.; Mahony, R.; Weiss, S. Equivariant filter design for inertial navigation systems with input measurement biases. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022; pp. 4333–4339.
22. van Goor, P.; Mahony, R. An equivariant filter for visual inertial odometry. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–5 June 2021; pp. 14432–14438.
23. Trawny, N.; Roumeliotis, S.I. *Indirect Kalman Filter for 3D Attitude Estimation*; Technical Report; University of Minnesota, Department of Computer Science and Engineering; Minneapolis, MN, USA, 2005; Volume 2.
24. Maybeck, P. *Stochastic Models, Estimation, and Control*; Academic Press: Cambridge, MA, USA, 1982.
25. Beder, C.; Steffen, R. Determining an initial image pair for fixing the scale of a 3d reconstruction from an image sequence. In Proceedings of the Joint Pattern Recognition Symposium, Berlin, Germany, 12–14 September 2006; pp. 657–666.
26. Eudes, A.; Lhuillier, M. Error propagations for local bundle adjustment. In Proceedings of the 2009 IEEE Conference On Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 2411–2418.
27. Hermann, R.; Krener, A. Nonlinear controllability and observability. *IEEE Trans. Autom. Control* **1977**, *22*, 728–740. [[CrossRef](#)]
28. Burri, M.; Nikolic, J.; Gohl, P.; Schneider, T.; Rehder, J.; Omari, S.; Achtelik, M.W.; Siegwart, R. The EuRoC micro aerial vehicle datasets. *Int. J. Robot. Res.* **2016**, *35*, 1157–1163. [[CrossRef](#)]
29. Campos, C.; Elvira, R.; Rodríguez, J.J.G.; Montiel, J.M.; Tardós, J.D. Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam. *IEEE Trans. Robot.* **2021**, *37*, 1874–1890. [[CrossRef](#)]
30. Grupp, M. Evo: Python Package for the Evaluation of Odometry and SLAM. Available online: <https://github.com/MichaelGrupp/evo> (accessed on 9 December 2024).
31. Li, M.; Mourikis, A.I. High-precision, consistent EKF-based visual-inertial odometry. *Int. J. Robot. Res.* **2013**, *32*, 690–711. [[CrossRef](#)]
32. Shuster, M.D. A survey of attitude representations. *Navigation* **1993**, *8*, 439–517.

33. Dam, E.B.; Koch, M.; Lillholm, M. *Quaternions, Interpolation and Animation*; Datalogisk Institut, Københavns Universitet: Copenhagen, Denmark, 1998; Volume 2.
34. Gelman, H. A note on the time dependence of the effective axis and angle of rotation. *J. Res. Natl. Bur. Stand.* **1971**, *75*, 165–171. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.