A Rotation-Translation-Decoupled Solution for Robust and Efficient Visual-Inertial Initialization Supplementary Material

Yijia He *
Chinese Academy of Sciences

Bo Xu *
Wuhan University

Zhanpeng Ouyang
ShanghaiTech University

Hongdong Li
Australian National University

1. Derivation of gyroscope bias optimizer

As introduced in the main paper Sec. 4.1, we describe how our optimizer calculate gyroscope bias over multiple frames. The core loss function is Eq. (7) in the main manuscript. The derivation of this loss function in detail will be described. As mentioned in Eq. (4) and Eq. (5) in the main manuscript, we obtain the relationship between the eigenvalues of $\mathbf{M}_{ij}$ and $\mathbf{R}_{c_ic_j}$ respectively, and we also get the transformation of $\mathbf{R}_{c_ic_j}$ and $\mathbf{R}_{b_ib_j}$, combine them:

$$\mathbf{M}_{ij} = \sum_{k=1}^{n} \left(\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R}_{bc}^\top \mathbf{R}_{b_ib_j} \mathbf{R}_{bc} \mathbf{f}_j^k\right) \left(\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R}_{bc}^\top \mathbf{R}_{b_ib_j} \mathbf{R}_{bc} \mathbf{f}_j^k\right)^\top \tag{1}$$

Further substitute Eq. (3) in the main manuscript into the above equation, the relationship between $\mathbf{M}_{ij}$ and $\mathbf{b}_g$ is derived, and we define $\mathbf{M}_{ij}$ related to $\mathbf{b}_g$ to be $\mathbf{M}'_{ij}$:

$$\mathbf{M}'_{ij} = \sum_{k=1}^{n} \left(\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R}_{bc}^\top \boldsymbol{\gamma}_{b_j}^{b_i} \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{R}_{bc} \mathbf{f}_j^k\right) \left(\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R}_{bc}^\top \boldsymbol{\gamma}_{b_j}^{b_i} \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{R}_{bc} \mathbf{f}_j^k\right)^\top \tag{2}$$

Define $\mathbf{R} = \mathbf{R}_{bc}^\top \boldsymbol{\gamma}_{b_j}^{b_i}$:

$$\mathbf{M}'_{ij} = \sum_{k=1}^{n} \left(\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R} \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{R}_{bc} \mathbf{f}_j^k\right) \left(\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R} \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{R}_{bc} \mathbf{f}_j^k\right)^\top \tag{3}$$

Because $\mathbf{M}'_{ij}$ contains the cross product of vector and rotation, we use the cross product property $\lfloor \mathbf{f}_i^k \rfloor_\times \mathbf{R} = \mathbf{R} \lfloor \mathbf{R}^\top \mathbf{f}_i^k \rfloor_\times$ to further simplify.

$$\mathbf{M}'_{ij} = \sum_{k=1}^{n} \left(\mathbf{R} \lfloor \mathbf{R}^\top \mathbf{f}_i^k \rfloor_\times \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{R}_{bc} \mathbf{f}_j^k\right) \left(\mathbf{R} \lfloor \mathbf{R}^\top \mathbf{f}_i^k \rfloor_\times \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{R}_{bc} \mathbf{f}_j^k\right)^\top \tag{4}$$

Finally, define $\mathbf{f}_i^{k\prime} = \mathbf{R}^\top \mathbf{f}_i^k$, and $\mathbf{f}_j^{k\prime} = \mathbf{R}_{bc} \mathbf{f}_j^k$. Eq. (7) in the main manuscript is derived.

$$\mathbf{M}'_{ij} = \sum_{k=1}^{n} \left(\mathbf{R} \lfloor \mathbf{f}_i^{k\prime} \rfloor_\times \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{f}_j^{k\prime}\right) \left(\mathbf{R} \lfloor \mathbf{f}_i^{k\prime} \rfloor_\times \operatorname{Exp}\left(\mathbf{J}_{\mathbf{b}_g}^{\gamma_{b_j}^{b_i}} \mathbf{b}_g\right) \mathbf{f}_j^{k\prime}\right)^\top \tag{5}$$

Because $\mathbf{R}$ is an orthogonal matrix and does not affect the eigenvalues of $\mathbf{M}'_{ij}$, $\mathbf{R}$ can be ignored.

*Equal contribution

2. Deviation of velocity and gravity estimator

2.1 Deviation of tightly-coupled solution

In Eq. (12) in the main manuscript, we use LiGT constraint [1] and the information of the visual-inertial system to get a constrained least squares problem which values to be estimated are velocity and gravity. In this section, we show the detailed derivation of this equation.

First, combine Eq. (9) and Eq. (5) in the main manuscript:

$$\mathbf{B}\mathbf{R}_{bc}^{\top}\left(\mathbf{p}_{b_1b_r}+\mathbf{R}_{b_1b_r}\mathbf{p}_{bc}-\mathbf{p}_{bc}\right)+\mathbf{C}\mathbf{R}_{bc}^{\top}\left(\mathbf{p}_{b_1b_i}+\mathbf{R}_{b_1b_i}\mathbf{p}_{bc}-\mathbf{p}_{bc}\right)+\mathbf{D}\mathbf{R}_{bc}^{\top}\left(\mathbf{p}_{b_1b_l}+\mathbf{R}_{b_1b_l}\mathbf{p}_{bc}-\mathbf{p}_{bc}\right)=0 \tag{6}$$

Define $\mathbf{B}'=\mathbf{B}\mathbf{R}_{bc}^{\top}, \mathbf{C}'=\mathbf{C}\mathbf{R}_{bc}^{\top}, \mathbf{D}'=\mathbf{D}\mathbf{R}_{bc}^{\top}, \mathbf{h}_m=\mathbf{R}_{b_1b_m}\mathbf{p}_{bc}-\mathbf{p}_{bc}$, we can get a simpler equation:

$$\mathbf{B}'\left(\mathbf{p}_{b_1b_r}+\mathbf{h}_r\right)+\mathbf{C}'\left(\mathbf{p}_{b_1b_i}+\mathbf{h}_i\right)+\mathbf{D}'\left(\mathbf{p}_{b_1b_l}+\mathbf{h}_l\right)=0 \tag{7}$$

In the above formula, there are global translations that can be used to connect the IMU information with the visual information, substituting into Eq. (1) in the main manuscript.

$$\begin{aligned}\mathbf{B}'\left(\mathbf{v}_{b_1}^{b_1}\Delta t_{1r}-\frac{1}{2}\mathbf{g}^{b_1}\Delta t_{1r}^2+\boldsymbol{\alpha}_{b_r}^{b_1}+\mathbf{h}_r\right)+\mathbf{C}'\left(\mathbf{v}_{b_1}^{b_1}\Delta t_{1i}-\frac{1}{2}\mathbf{g}^{b_1}\Delta t_{1j}^2+\boldsymbol{\alpha}_{b_i}^{b_1}+\mathbf{h}_i\right)+\\ \mathbf{D}'\left(\mathbf{v}_{b_1}^{b_1}\Delta t_{1l}-\frac{1}{2}\mathbf{g}^{b_1}\Delta t_{1l}^2+\boldsymbol{\alpha}_{b_l}^{b_1}+\mathbf{h}_l\right)=0\end{aligned} \tag{8}$$

Let $\mathbf{s}_{1m}=\boldsymbol{\alpha}_{b_m}^{b_1}+\mathbf{h}_m=\boldsymbol{\alpha}_{b_m}^{b_1}+\mathbf{R}_{b_1b_m}\mathbf{p}_{bc}-\mathbf{p}_{bc}$, where $m\in r,i,l$. The system equation is then given by:

$$\begin{aligned}&\mathbf{B}'\left(\mathbf{v}_{b_1}^{b_1}\Delta t_{1r}-\frac{1}{2}\mathbf{g}^{b_1}\Delta t_{1r}^2\right)+\mathbf{C}'\left(\mathbf{v}_{b_1}^{b_1}\Delta t_{1i}-\frac{1}{2}\mathbf{g}^{b_1}\Delta t_{1j}^2\right)+\mathbf{D}'\left(\mathbf{v}_{b_1}^{b_1}\Delta t_{1l}-\frac{1}{2}\mathbf{g}^{b_1}\Delta t_{1l}^2\right)=-\mathbf{B}'\mathbf{s}_{1r}-\mathbf{C}'\mathbf{s}_{1i}-\mathbf{D}'\mathbf{s}_{1l}\\ &\left(\mathbf{B}'\Delta t_{1r}+\mathbf{C}'\Delta t_{1i}+\mathbf{D}'\Delta t_{1l}\right)\mathbf{v}_{b_1}^{b_1}+\frac{1}{2}\left(\mathbf{B}'\Delta t_{1r}^2+\mathbf{C}'\Delta t_{1i}^2+\mathbf{D}'\Delta t_{1l}^2\right)\mathbf{g}^{b_1}=-\mathbf{B}'\mathbf{s}_{1r}-\mathbf{C}'\mathbf{s}_{1i}-\mathbf{D}'\mathbf{s}_{1l}\end{aligned} \tag{9}$$

Finally, we get the equation form consistent with the main manuscript:

$$\mathbf{A}_1\mathbf{v}_{b_1}^{b_1}+\mathbf{A}_2\mathbf{g}^{b_1}=\mathbf{d} \tag{10}$$

where

$$\begin{aligned}\mathbf{A}_1&=\mathbf{B}'\Delta t_{1r}+\mathbf{C}'\Delta t_{1i}+\mathbf{D}'\Delta t_{1l}\\ \mathbf{A}_2&=\frac{1}{2}\left(\mathbf{B}'\Delta t_{1r}^2+\mathbf{C}'\Delta t_{1i}^2+\mathbf{D}'\Delta t_{1l}^2\right)\\ \mathbf{d}&=-\mathbf{B}'\mathbf{s}_{1r}-\mathbf{C}'\mathbf{s}_{1i}-\mathbf{D}'\mathbf{s}_{1l}\end{aligned} \tag{11}$$

The initial velocity and gravity direction variables can be solved with Eq. (10).

2.2 Deviation of Loosely-coupled solution

In the main paper, we introduced how to use the IMU model to establish the residual to solve multiple frames of velocity, gravity, and scale under the premise of known positions. Here, we will deduce in detail how to use the IMU model to establish the residual. The translation and velocity part of IMU pre-integration observation model are defined same as main manuscript, as follows:

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_{b_k}^{b_i}&=\mathbf{R}_{b_ic_0}\left(s\left(\mathbf{p}_{c_0b_k}-\mathbf{p}_{c_0b_i}\right)+\frac{1}{2}\mathbf{g}^{c_0}\Delta t_{ik}^2\right)-\mathbf{v}_{b_i}^{b_i}\Delta t_{ik}\\ \hat{\boldsymbol{\beta}}_{b_k}^{b_i}&=\mathbf{R}_{b_ic_0}\left(\mathbf{R}_{c_0b_k}\mathbf{v}_{b_k}^{b_k}+\mathbf{g}^{c_0}\Delta t_{ik}\right)-\mathbf{v}_{b_i}^{b_i}\end{aligned} \tag{12}$$

Let $s\mathbf{p}_{c_0b_k}=s\mathbf{p}_{c_0c_k}-\mathbf{R}_{c_0b_k}\mathbf{p}_{bc}$, and substitute it into Eq. (12):

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_{b_k}^{b_i}&=\mathbf{R}_{b_ic_0}\left(\left(s\mathbf{p}_{c_0c_k}-\mathbf{R}_{c_0b_k}\mathbf{p}_{bc}-\left(s\mathbf{p}_{c_0c_i}-\mathbf{R}_{c_0b_i}\mathbf{p}_{bc}\right)\right)+\frac{1}{2}\mathbf{g}^{c_0}\Delta t_{ik}^2\right)-\mathbf{v}_{b_i}^{b_i}\Delta t_{ik}\\ \hat{\boldsymbol{\beta}}_{b_k}^{b_i}&=\mathbf{R}_{b_ic_0}\left(\mathbf{R}_{c_0b_k}\mathbf{v}_{b_k}^{b_k}+\mathbf{g}^{c_0}\Delta t_{ik}\right)-\mathbf{v}_{b_i}^{b_i}\end{aligned} \tag{13}$$

In the above formula, only $\mathbf{v}_{b_i}^{b_i}$, $\mathbf{v}_{b_k}^{b_k}$ and $\mathbf{g}^{c_0}$ are unknowns. Eq. (17) in the main manuscript can be derived:

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_{b_k}^{b_i} &= s\mathbf{R}_{b_i c_0}\left(\mathbf{p}_{c_0 c_k} - \mathbf{p}_{c_0 c_i}\right) - \mathbf{R}_{b_i c_0}\mathbf{R}_{c_0 b_k}\mathbf{p}_{bc} + \mathbf{p}_{bc} + \frac{1}{2}\mathbf{R}_{b_i c_0}\mathbf{g}^{c_0}\Delta t_{ik}^2 - \mathbf{v}_{b_i}^{b_i}\Delta t_{ik} \\ \hat{\boldsymbol{\beta}}_{b_k}^{b_i} &= \mathbf{R}_{b_i c_0}\mathbf{R}_{c_0 b_k}\mathbf{v}_{b_k}^{b_k} + \mathbf{R}_{b_i c_0}\mathbf{g}^{c_0}\Delta t_{ik} - \mathbf{v}_{b_i}^{b_i} \end{aligned} \tag{14}$$

Finally, use Eq. (18) in the main manuscript and arrange to obtain:

$$\begin{aligned} \mathbf{H}' &= \begin{bmatrix} -\mathbf{I}\Delta t_{ik} & \mathbf{0} & \mathbf{R}_{b_i c_0}\left(\mathbf{p}_{c_0 c_k} - \mathbf{p}_{c_0 c_i}\right) & \frac{1}{2}\mathbf{R}_{b_i c_0}\Delta t_{ik}^2 \\ -\mathbf{I} & \mathbf{R}_{b_i c_0}\mathbf{R}_{c_0 b_k} & \mathbf{0} & \mathbf{R}_{b_i c_0}\Delta t_{ik} \end{bmatrix} \\ \mathbf{b}' &= \begin{bmatrix} \boldsymbol{\alpha}_{b_k}^{b_i} - \mathbf{p}_{bc} + \mathbf{R}_{b_i c_0}\mathbf{R}_{c_0 b_k}\mathbf{p}_{bc} \\ \boldsymbol{\beta}_{b_k}^{b_i} \end{bmatrix} \end{aligned} \tag{15}$$

3. Additional results for accuracy evaluation

We note that we have reported the accuracy comparison of different algorithms on the EuRoC dataset classified by angular velocity in the main paper Tab.1. However, previous work [2–4] usually counts the results of the 11 sequences on the EuRoC dataset separately instead of classifying them according to the angular velocity. In order to maintain consistency, we also present the results of the 11 sequences in this section. The same experimental strategy and evaluation method from the main paper is adopted.

As shown in Tab.1, some conclusions consistent with the main paper can be obtained. DRT-l still outperforms other initialization methods on almost all the sequences, which demonstrates the effectiveness of the proposed rotation-translation decoupled method. It is worth noting that due to the rapid movement on the V103 and V203 sequences, the number of data segments that can be successfully initialized by AS-MLE and VINS-Mono is too few to be statistically significant. Finally, the gravity vector error of all methods exceeds 3 degrees on the V101 sequence, due to the accelerometer bias of about $0.5m/s^2$ for this dataset. All methods assume that the accelerometer bias is 0 except AS-MLE which introduces a prior constraint to estimate it. We re-evaluated with VINS-Mono and DRT-t after subtracting a bias of 0.5 from the acceleration, and the gravity error has dropped from 3.21 to 0.91 and 3.18 to 0.70 respectively.

Table 1. Exhaustive initialization results for 10KFs setting in per dataset from EuRoC. For each metric, the best in **red**, the second best in blue, - means that the effective initialization segments are less than 10, which is considered as failure

	Dataset	MH01	MH02	MH03	MH04	MH05	V101	V102	V103	V201	V202	V203
Scale RMSE	**AS-MLE**	0.24	0.21	0.34	0.37	0.46	0.24	0.41	-	0.31	0.24	-
	CS-VISfM	0.53	0.50	0.48	0.57	0.55	0.57	0.48	0.33	0.49	0.43	0.31
	CS-VISfM-GBE	0.25	0.21	0.15	0.31	0.31	0.18	0.07	0.12	0.23	0.08	**0.09**
	VINS-Mono	0.18	0.18	0.19	**0.20**	**0.27**	0.15	0.15	-	0.25	0.15	-
	DRT-t	0.27	0.21	0.14	0.31	0.36	0.19	0.06	**0.11**	0.21	0.09	0.10
	DRT-l	**0.13**	**0.13**	**0.11**	0.21	**0.27**	**0.12**	**0.04**	0.27	**0.16**	**0.06**	0.28
	DRT-l-wo-GBE	0.49	0.47	0.47	0.55	0.46	0.40	0.50	0.39	0.52	0.46	0.40
velocity RMSE	**AS-MLE**	0.11	0.09	0.30	0.25	0.25	0.07	0.24	-	0.08	0.13	-
	CS-VISfM	0.20	0.20	0.34	0.47	0.24	0.15	0.27	0.22	0.15	0.20	0.26
	CS-VISfM-GBE	0.10	0.09	0.14	0.22	0.20	0.08	0.07	**0.08**	0.08	0.05	**0.09**
	VINS-Mono	0.08	0.08	0.19	**0.14**	**0.17**	0.06	0.11	-	0.07	0.08	-
	DRT-t	0.11	0.11	0.14	0.19	0.22	0.07	0.06	0.11	0.08	0.06	**0.09**
	DRT-l	**0.07**	**0.06**	**0.11**	**0.14**	**0.17**	**0.05**	**0.04**	0.21	**0.05**	**0.04**	0.15
	DRT-l-wo-GBE	0.21	0.17	0.37	0.30	0.26	0.16	0.25	0.20	0.13	0.21	0.41
G.Dir RMSE	**AS-MLE**	1.64	1.36	3.06	1.87	2.61	3.37	5.57	-	1.79	2.68	-
	CS-VISfM	5.76	5.68	6.00	5.68	6.32	8.23	5.22	5.60	6.13	5.23	5.92
	CS-VISfM-GBE	**0.94**	**0.92**	**0.90**	0.99	0.98	3.20	**0.79**	**1.39**	1.12	**0.79**	**1.05**
	VINS-Mono	1.17	1.11	1.57	1.23	1.42	3.21	2.60	-	1.43	1.26	-
	DRT-t	1.00	0.97	0.96	0.99	0.99	**3.18**	**0.79**	1.48	1.15	0.84	1.17
	DRT-l	0.95	0.94	0.93	**0.94**	**0.92**	3.22	0.85	2.08	**1.10**	0.88	1.28
	DRT-l-wo-GBE	5.68	5.50	5.39	6.20	5.92	8.01	4.98	5.46	5.95	5.34	7.89

References

[1] Qi Cai, Lilian Zhang, Yuanxin Wu, Wenxian Yu, and Dewen Hu. A pose-only solution to visual reconstruction and navigation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. 2

[2] Carlos Campos, José MM Montiel, and Juan D Tardós. Inertial-only optimization for visual-inertial initialization. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 51–57. IEEE, 2020. 3

[3] Yunwen Zhou, Abhishek Kar, Eric Turner, Adarsh Kowdle, Chao X Guo, Ryan C DuToit, and Konstantine Tsotsos. Learned monocular depth priors in visual-inertial initialization. *arXiv preprint arXiv:2204.09171*, 2022. 3

[4] David Zuñiga-Noël, Francisco-Angel Moreno, and Javier Gonzalez-Jimenez. An analytical solution to the imu initialization problem for visual-inertial systems. *IEEE Robotics and Automation Letters*, 6(3):6116–6122, 2021. 3