Simultaneous State Initialization and Gyroscope Bias Calibration in Visual Inertial aided Navigation

Responses to Reviewer Comments

Jacques Kaiser, Agostino Martinelli, Flavio Fontana and Davide Scaramuzza

We would like to thank the editor and the reviewers for the time they spent reading and carefully reviewing the manuscript, and for their constructive suggestions, which helped us to improve the paper. Below, is our response to the reviewers' remarks with the changes that we've made to the document.

Reviewer #1:
The authors study empirically the impact of noise on a recently derived closed-form solution [1] that determines the angular and linear velocity of a moving platform by fusing visual and inertial measurements. The main contributions, as claimed by the authors, are twofold: (a) gyro bias is (empirically) shown to have significant impact on performance; (b) a new method is introduced for online estimation of the gyro bias.

Overall, the paper is well-written, tries to be self-contained, and is easy to follow. However, in my opinion, it is not sufficiently innovative and lacks thorough analysis to meet the high-standards of RA-L. On the other hand, the topic is important and the presented results and proposed improvement are impressive, and thus are of interest to the robotics community. Therefore, the paper is suitable for publication in ICRA.

Detailed comments follow below.

The authors generally do a good job addressing related work in visual-inertial aided navigation and visual-inertial structure from motion, mentioning both filtering and optimization based approaches. I would suggest to add also the following related publications:

- Huang, Guoquan, Michael Kaess, and John J. Leonard. "Towards consistent visual-inertial navigation." In Robotics and Automation (ICRA), 2014 IEEE International Conference on, pp. 4926-4933. IEEE, 2014.

- Indelman, Vadim, Stephen Williams, Michael Kaess, and Frank Dellaert. "Information fusion in navigation systems via factor graph based incremental smoothing." Robotics and Autonomous Systems 61, no. 8 (2013): 721-738.

- Hesch, Joel, Dimitrios G. Kottas, Sean L. Bowman, and Stergios Roumeliotis. "Consistency analysis and improvement of vision-aided inertial navigation." Robotics, IEEE Transactions on 30, no. 1 (2014): 158-176.

Answer #1.1: These references have been included.

The paper builds upon [1] and shows empirically, using both simulated and real-world data, that gyroscope bias has significant impact of performance of the closed-form solution of [1]. My main concern, however, is that in its current version the study is purely empirical and does not attempt to investigate sensitivity of the closed-form solution using more analytical tools. This would shed further light, I believe, to the observation made by the authors, and allow addressing my next comment.

Answer #1.2: We agree that most of the results have been obtained empirically. On the other hand, obtaining analytic results that describe the sensitivity of the performance on the gyro Bias seems prohibitive. The effect of the gyro bias in (1) is very troublesome since the spurious rotation due to the bias cannot be separated from the true rotation and this because rotations are in general non-commutative. Additionally, even if true rotations do not occur, the change of the left-hand side in (1) due to the bias, is non-trivial for a generic acceleration. Similarly, providing analytic properties of the cost function, seems to be prohibitive since it depends on the entire trajectory. See also the next answer #1.4

One source of confusion is that the impact of gyroscope bias on performance was already considered in Monte-Carlo study in [1] (section 5.2). This fact should be mentioned in the present paper and the difference between the two studies should be discussed. In particular, it seems that the conclusion regarding the significant impact of gyroscope bias on performance should have been reached in both cases; however that appears to be not the case. A discussion of these aspects would be beneficial, in my opinion.

Answer #1.3: We explicitly mentioned the differences with [1] at the beginning of section IV and in the last paragraph of IV.D.

Another point that deserves clarification is the way the cost function in eq. 4 is minimized. Since that optimization process requires initialization, can it be stuck in local minima? An intuitive explanation why the symmetry in the cost function is a bad thing would be also beneficial (Section V-B).

Answer #1.4: It is true. This can occur in principle. Extensive simulations clearly show that the function is locally convex (i.e., in the region around the true bias). This has now been mentioned in section V.A. If the time of integration is small, and the rotations only occur around a single axis, this local convexity is lost. In particular, the cost function exhibits a symmetry: it is invariant with respect to the component of the bias collinear with the gravity. An explanation why the symmetry in the cost function is a bad thing has been added in the last paragraph of V.A.

Minor comment - please indicate where [19] was published.

Answer #1.5: Done

Reviewer #2:
This paper presents an algorithm which obtains closed-form estimates for the initial values for the velocity and gravity vectors and sensor biases (along with feature depths) for the problem of visual-inertial localization. The algorithm is largely a modification of an existing method, in which case the original closed-form solution was detailed.

The paper begins by briefly explaining the method in Ref. 1. Of note is the specific emphasis on Eq. 10 in the original paper (Eq. 3 in the current paper) where the gravity vector magnitude is constrained.

Sec. IV details limitations which the authors point out exist in the original method. To show this, a combination of real and synthetic data is used, although it is not clear why this is the case. It is mentioned that the justification for using synthetic camera data is the availability of ground-truth landmark depth values, for comparison. However, it seems that by that logic it would also make sense to use simulated pose and measurement data as well. This renders Figure 3 rather confusing. Presumably the relative error axis is compared with ground-truth data. However since the paper implies that real data was used for position and inertial measurements, what was the source of the ground truth data for these terms?

Answer #2.1: Section IV has totally been re-written. Now all the measurements are synthetic, as suggested.

Furthermore, it is unclear why the comparison in Fig. 3 is made. It seems as though the figure serves to show that the gravity refinement outlined in the original algorithm does not significantly improve the result of the experiment. Is this used as a basis to remove this constraint in the following sections? If so, please explicitly state this. Its also recommended to outline or discuss any cases where the authors feel that constraining the magnitude of the gravity vector could be beneficial.

Answer #2.2: Now we clearly state at the end of section IV-B that we remove the gravity constraint after this section. Since the gravity is well-estimated all the time, we do not feel that constraining its magnitude could be beneficial.

Sec. IV.C states that accelerometer bias (within the attempted ranges) does not seriously undermine the original algorithm. The authors also state that the variant of the original algorithm which estimates the biases is not robust in the experiment due to insignificant effect. It is recommended to expand on this as it is unclear. Perhaps a comparison between the two variants? Does the situation change if the biases are large?

Answer #2.3: Actually, the bias on the accelerometer cannot be estimated for a motion where important rotations, at least around two independent axes, do not occur. Indeed, the accelerometer bias is non-observable if these rotations do not occur (see property 12 and/or table 2 in [1]). This has been mentioned in a footnote of section IV.C. Additionally, in the first paragraph of section IV.C, we provide an explanation for the negligible effect of the accelerometer bias.

Finally it is shown that the gyroscope bias has a significant impact on the solution. This reviewer feels that this may have more significance if all the data were generated via simulation. The use of mixed data (with inherent biases and errors) potentially undermines the precise significance of a particular source of added error.

Answer #2.4: We fully agree with this and we followed this suggestion.

Sec. V outlines the modifications made to the original algorithm to estimate the gyro bias. The specific formulation is unclear to this reviewer. Eq. 4 changes the original closed-form solution to an optimization. It is not clear what form the original matrices take when computed with respect to B. The paper hints that Eq. 4 is solved via an optimization. Is the solution still closed form? Or do gradients need to be computed to optimize the non-linear constraints? Please consider adding detail to this section, as in this reviewer's opinion it constitutes the majority of the theoretical contribution of the work.

If the closed-form solution has indeed been turned into an iterative optimization, what are the implications for performance? Please consider adding performance metrics to the paper, as initialization is a time-sensitive stage of visual-inertial estimation.

Answer #2.5: See answer #1.4

Please consider adding a comparison metric to Figure 6. It is not clear if the values that the solution converged to are correct. Were these bias values the ones that were used to corrupt the measurements? What about the pre-existing bias in those measurements since they were from real sensor data?

Answer #2.6: Now that we perform our evaluations on purely synthetic measurements, we can precisely provide the ground truth bias. Moreover, the relative error between the real bias and the bias estimate has been added.

Sec. V.C outlines a certain symmetry in the cost function. This problem is solved by adding a regularizer which penalizes large bias estimates. To this reviewer, this does not seem like a principled solution. It penalizes all gyro bias axes equally it is possible to know which axis is continually collinear with gravity, and regularize only that axis. This regularizer could also potentially result in erroneous estimates if sufficient excitation did exist during initialization and the correct bias values were in fact very large.

If the regularizer is used to include prior information, it seems that the value of lambda should be set given the covariance of these prior estimates, rather than empirically.

Answer #2.7: We really thank the reviewer for this remark. The cost function has been modified accordingly (see equation (5)). In the last paragraph of section V.C we also added the remark about the regularizer.

Section VI describes experiments using real data, where it is unclear (in Fig. 11) whether there is any significant advantage in using the modified algorithm. As the authors themselves state, this may be because the estimated bias values are very small. The merits of the proposed solution as opposed to the original algorithm would be more convincing if an experiment was presented where the results were significantly improved.

Answer #2.8: We added two new figures where we corrupted the true measurements with two artificial biases. In fig 9b the artificial bas is 0.05 rad/s in fig 9c is 0.1 rad/sec. Now, the optimized solution significantly outperforms the original solution. This because the new solution estimated very accurately the bias and hence the performance is unaffected by this artificial bias.

Editor:
In this paper, the authors consider the problem of extending the algorithm in [1] for initialisation to consider real data and, in particular, the impact of gyro biases. To address the issue with the gyro biases, a regularisation scheme is introduced. Both reviewers feel that the paper is well-written and interesting. However, they have the following concerns:

1. Both reviewers express concerns about the lack of innovation in the paper. R1 notes that the analysis is almost entirely empirical and lacks any theoretical analysis. R2 argues that the work builds upon [1] by adding a nonlinear optimization scheme and regulariser to compute the gyro biases. However, there are many questions about the validity and performance of this approach.

Answer #E.1: *See answers #1.2, #1.4, and #2.7*

2. R2 expresses the additional concern that the results mix both real data and simulations because this can mix in unmodelled effects. This raises the question of why the analysis in Section IV did not simply use the output from a high-quality simulator.

Answer #E.2: *See answers #2.1, #2.4*

3. R1 expressed the concern that the paper did not fully describe the study in section 5.2 of [1]. Checking this, the reviewer is correct in that the study does include time varying biases on both the accelerometers and the gyros jointly - results for each independently of one another do not appear to be presented. This clarification is unclear in the paper.

Answer #E.3: *See answer #1.3*

4. R2 expressed concerns about the discussion in Section IV.C and biases. The discussion in the section is unclear about the similarities and differences between the algorithms and [1]. To me, it raises the question of whether the accelerometer readings actually play a significant role with the algorithms.

As regards the relative contribution of the biases, I think it's reasonably clear why the example, in Fig. 4 and Fig. 5, would produce very different-looking results. 0.2 rad/s is 11 deg/s drift, which is clearly very large. $0.2m/s^2$ *error leads, in my calculations to a worst case estimated error in orientation of about 0.8 degrees (atan2(0.2/sqrt(2),9.81-0.2/sqrt(2))*180/pi).*

Answer #E.4: *See answer #2.3*