Tightly Coupled Inertial Visual GNSS Solution - Application to LIDAR Mapping in Harsh and Denied GNSS Conditions

Pierre Bénet, Mourad Saidani, Alexis Guinamard
SBG Systems, Carrière sur Seine, France

BIOGRAPHY

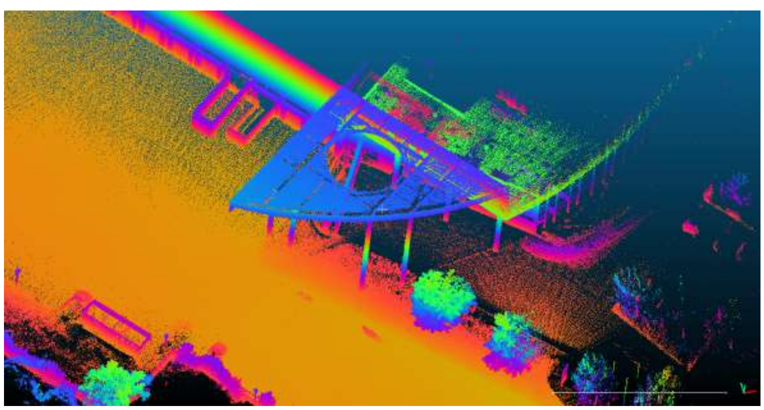
Pierre Bénet is a research engineer at SBG Systems since 2019. His research interests include Visual-Inertial Navigation for high precision localization and mapping.

Mourad Saidani is Inertial Navigation Algorithms Manager. He joined SBG in 2017. His research interests include GNSS and inertial coupling.

Alexis Guinamard is Chief Technology Officer at SBG Systems. He co-founded SBG Systems in 2007. His research interests include GNSS, Inertial sensors and inertial filtering and more lately vision.

ABSTRACT

Global navigation satellite system real-time kinematic (GNSS-RTK) positioning is today a key technology for survey and mapping applications. To extend the capability of GNSS in difficult environment, a tight coupling between GNSS-RTK and an inertial navigation systems can greatly improve the results. This solution is adapted to small GNSS outage under bridges and in urban canyon for automotive survey for instance. If the time spent in GNSS outage is too long or if the kinematic of the survey is too weak, the GNSS inertial solution can be compromised, due to high navigation errors, and ultimately, impossibility to align heading angle at initialization. This occurs most often for pedestrian survey where the GNSS conditions are worst and dynamics are low. Some solutions on the market propose a LIDAR based SLAM to overcome the limitations of INS/GNSS. However this solution has a great impact on the mapping solution in terms of bill of materials costs, and power consumption. Indeed, two LIDARs are generally used in this case : one dedicated to SLAM and one dedicated to point cloud generation.



This paper presents an innovative solution to overcome the GNSS/INS limitations, whereas minimizing the system complexity by using a tightly coupled GNSS/INS solution, coupled with our monocular visual inertial SLAM system (DVM). This solution is capable of initialization in a few seconds, and is very reliable in the long term. This vision/INS/GNSS coupling increases the overall RTK fix rate and broadens the availability of high precision navigation solutions under challenging conditions. In addition, our visual SLAM system can optimize the full visual graph and achieve cm accurate positioning on the full path in indoor, benefiting from GNSS points at the entrance and exit of the indoor survey, as well as visual loop closure. For more flexibility and accuracy, our visual graph optimizer can estimate the intrinsic and extrinsic calibration parameters of the camera using only a few GNSS points, allowing easy third-party camera aiding. Finally, visual inertial SLAM post-processing proposes an alternative to LIDAR SLAM that does not suffer from poor geometry issues. Going further we will assess the performance of our inertial visual GNSS solution by generating a LIDAR point cloud and analyzing the consistency of the point cloud.

I. INTRODUCTION

Nowadays, the LIDAR (Light Detection And Ranging) has become an indispensable technology for high precision mapping. The applications for LIDAR mapping are numerous, from architecture and cultural heritage to forestry and agriculture. For a complete reconstruction of an area to map, the LIDAR must often be deployed in areas only accessible by the pedestrian. This leads to the emergence of backpack solutions. Whereas point cloud density and map completeness increases, the accuracy of the point cloud has to deal with the difficult motion profile of the pedestrian and more difficult GNSS conditions. Typical scenario of a backpack survey is a continuous outdoor and indoor mapping of a building. The survey lasts from thirty minutes to several hours. The surveyor can also use a kick scooter or a Segway to map a larger outdoor area more easily.

The surveyor has often access to a snapshot of the map in real time to check if all the area has been correctly mapped. However, assessing the quality of the point cloud in real time is more difficult. Typical metric for assessing the quality of the point cloud is the ”best fitting-plane distance”. However, this metric does not give the precision of the full model [1]. The precision of the full model is mostly defined by the GNSS coverage during the survey. Different solutions are available to the surveyor to have a good GNSS coverage:

- If the surveyor wants to quickly process its acquisition, he needs an RTK setup with a close enough base station.
- If the surveyor can wait a little, a third party GNSS/INS post-processing software such as Qinertia can process the PPK.
- If the area to map is too far from a base station, the trajectory can still be post-processed using a Virtual Base Station system that Qinertia also provides [2]. This technology allows arbitrary large survey with continuous global position.

II. RELATED WORK

1. Backpack Mapping

Several commercial Lidar backpack systems have been developed in the past 10 years [1]. Hand-held systems usually do not support GNSS, whereas most backpack system support GNSS. Some systems only work in GNSS covered areas. A few backpacks use Dual GNSS antennas. As the two antennas need to be separated enough, the system becomes a little cumbersome and the dual antennas are removed before entering a building [1]. During large GNSS gaps, the trajectory is usually recovered using LIDAR SLAM. In our case, we use visual inertial SLAM. With the difficult motion profile of the pedestrian, the navigation initialization is a challenge itself. It is usually done by comparing the trajectory generated by a local SLAM with the GNSS points. As SLAM is able to yield a local consistent trajectory, it can be aligned by a few GNSS points at start using Singular Value Decomposition (SVD).

2. Visual Inertial GNSS Slam

a). Visual Inertial SLAM

An overview of visual inertial SLAM and detailed descriptions of high performance visual inertial SLAM systems can be found on ORB-SLAM3 [3] and on DVM [4]. Lately [5] pointed out several problems on the KITTI odometry dataset [6]. Besides the algorithms [5] developed to solve these problems, it underlines the importance of the sensor setup. Hence great care should be taken for the rigidity of the setup if stereo cameras are used [5]. Fixed focal length should be used [7]. Raw images should be available [5] [4] [3].

b). Loose Visual Inertial GNSS Coupling

Loose coupling between visual-inertial-SLAM and GNSS allows to obtain a globally positioned trajectory. Since SLAM is very reliable, it can afford to take only the accurate and robust GNSS points. Hence a first improvement for GNSS loose coupling is to reject all non RTK-fix GNSS points.

c). Tight Visual Inertial GNSS Coupling

However, sometimes the system operates in such a harsh environment that the SLAM system would benefit from additional satellite fixes. To do so without robustness issue, it becomes necessary to constrain the GNSS solution beforehand. This can be done with tight coupling. In tight coupling the GNSS solution will be sought in an area defined by the filter prediction. Then two paradigms are available:

- The optimal formulation adds the ambiguity states to the filter. GVINS [8] based on VINS Fusion [9] demonstrates tight coupling using 1, 2, and 3 satellites only.
- The federated filter feeds the GNSS estimator with the current filter prediction. This filter is sub-optimal, but it is enough to improve the solution over a loosely coupled strategy [10] [11] [12]

Our strategy will be similar to the federated filter.

III. VISUAL INERTIAL GNSS PROCESSING

1. Visual SLAM

The role of SLAM is to find features and localize the images for later global optimization. The features that are associated during SLAM are the same that are used for global optimization. Likewise the SLAM trajectory is a starting point of the global optimization. The SLAM is therefore a critical part for the accuracy of the post-processed solution. We use DVM [4], a Sliding Window Filter (SWF).

SWF are nowadays the most powerful methods for visual odometry. The filter behaves at the same time like a nonlinear optimization and an extended Kalman filter. This allows to use all information available at a time, having always the best linearization point possible. It works as follow: On every images, features are detected on high gradient location. Points of interest are associated on up to 10 other images using normalized sum of squared difference. About 1000 points are detected on a 1Mp image. The first 10 images are processed using gyrometer aided visual SLAM. Then, a local optimization is done to find accelerometer bias, gravity direction, speed and scale. Finally the inertial tight coupling begins. The sliding window system is composed of 6 keyframes that are continuously optimized and 4 fixed keyframes that are chosen among old keyframes allowing opportunistic loop closing. This loop closing strategy can be classified as mid-term data association [3] allowing zero-drift in already mapped areas.

As a critical part of the workflow, we improved the initialization robustness of our SLAM system by introducing a smooth transition between the gyrometer aided SLAM and the inertial tight coupling SLAM. We also improved the long term robustness by preventing the inertial biases to drift too much. The assumption is that the in run biases should never exceed the initial biases uncertainty. To model this phenomenon, we use a correlated noise to model the inertial biases. The rate random walk is unchanged but the correlation time is tuned so that the peak of the correlated noise on the Allan variance equals the initial bias uncertainty.

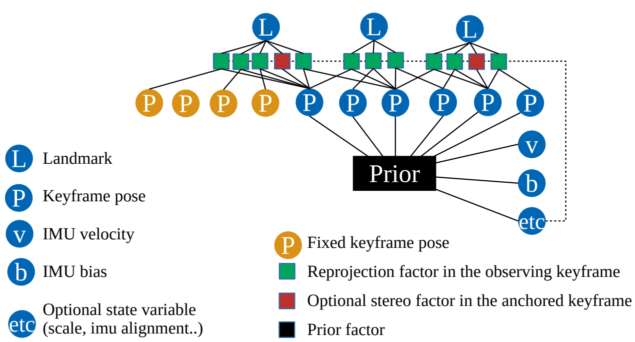


Figure 1: Structure of our SLAM system with 10 keyframes among them 4 are fixed. At each local bundle adjustment the state variables of the landmark (their inverse depth) are subtituted using the Schur Complement to obtain a small and dense information matrix.

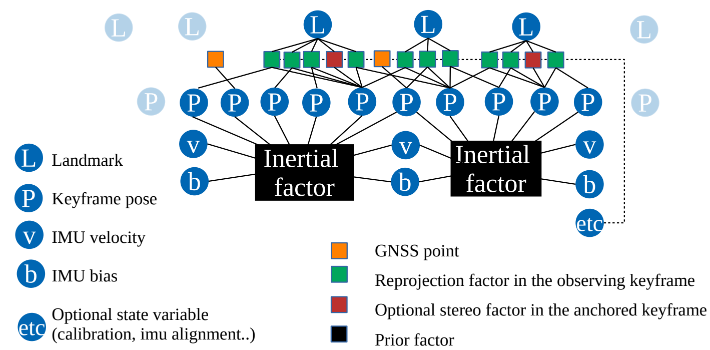


Figure 2: Structure of our global bundle adjustment system. At each global bundle adjustment the state variables of the landmarks are substituted using the Schur Complement to form a big and sparse information matrix. Resolution is done using a conjugate gradient with block-diagonal preconditioning. In order to reduce the problem size we use only 1 inertial state every 5 visual states, assuming that the inertial preintegration does not drift too much after 5 images.

2. Generalized IMU Preintegration

One of our main contribution is visual inertial global bundle adjustment with generalized IMU preintegration. That means that the inertial postprocessing is not limited to basic inertial states such as accelerometer bias and gyrometer bias but it can easily be extended to take into account more errors such as gyrometer scale factor and accelerometer scale factor for example. It is also capable of camera calibration which is a typical feature of photogrammetry software.

For global bundle adjustment, a sparse non linear least square is build containing the whole trajectories and variables to estimate. As it optimizes the full trajectory at onces, it is incompatible with the usual Extended Kalman Filter for inertial Filtering which works in a sequential manner. Imu preintegration have recently been successfully used in visual inertial local bundle adjustment [13] [3] but it is limited to simple IMU model and one inertial parameters per images. In our case, we use a general IMU model and fewer inertial parameters than images, yielding convergence speed up and robustness improvements (see Figure 2). To reach that goal we pose the inertial factor as a function of a local inertial state that predicts a few positions and attitudes ahead and the next inertial state:

$$h(p_k, b_k, v_k, p_{k+1} ... p_{k+5}, b_{k+5}, v_{k+5}) = g(p_k, b_k, v_k) - (p_k ... p_{k+5}, b_{k+5}, v_{k+5}) \quad (1)$$

with p_k the pose at time k, b_k the inertial biases at time k and v_k the speed at time k.

To use it in the global bundle adjustment, the covariance, gradient and value of h needs to be computed. This is possible by reusing the jacobian and propagation procedures of a conventional Extended Kalman Filter [14]. Hence g in equation 1 and its gradient is computed by stacking successive EKF propagation results from a classic EKF prediction function f [15].

$$g(x_k) = \begin{matrix} x_k + \int_k^{k+1} f(x_t)dt \\ ... \\ x_k + \int_k^{k+5} f(x_t)dt \end{matrix} \quad (2)$$

$$G(x_k) = \begin{matrix} \Pi_k^{k+1}(I + \frac{\partial f}{\partial x}(x_t)dt) \\ ... \\ \Pi_k^{k+5}(I + \frac{\partial f}{\partial x}(x_t)dt) \end{matrix} \quad (3)$$

The covariance of g is recovered the same way the prior is computed in figure 1 using the variable state dimension filter theory [16] [4]. Its inverse will give the weight of the observation in the global bundle adjustment. Once these quantity have been computed, the generated information matrix and error must be carefully added at the right indexes to the global sparse system.

3. GNSS Tight Coupling

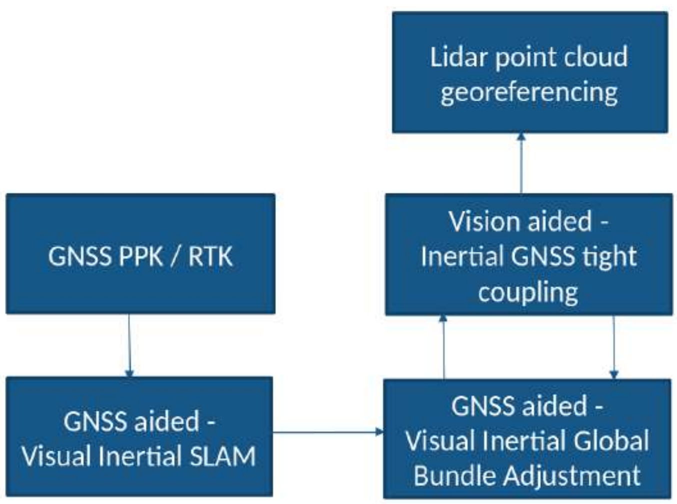


Figure 3: Scheme of the processing workflow in case of difficult GNSS conditions.

The post processing alternated between GNSS aided visual inertial global BA and visual SLAM aided PPK. Thus the position is iteratively refined. Feature position and imu biases are improving at the same time as GNSS ambiguity are resolved and refined. The inertial GNSS tight coupling is done by a modified version of SBG Qinertia software that is capable of taking the positionning of the visual SLAM as an additive aiding.

4. Performance Assessment

a). Difficult GNSS Condition

This test focuses on increasing the available RTK fix rate. This experience shows the efficiency of a tight coupling strategy over a loose coupling strategy. We did a full size test of 30 minutes (see Figure 4). The backpack went in an old town with narrow streets, passages with stairs, where townhouse are often blocking the sky. But the most difficult part was done along a block of flats with balconies, this corresponds to the bottom left corner of the trajectory. Analysis of the solution over a classic GNSS-inertial tight coupling in figure 5 shows that vision helps to get the strategic GNSS fixes that are necessary to have a full centimetric trajectory in such harsh GNSS conditions.

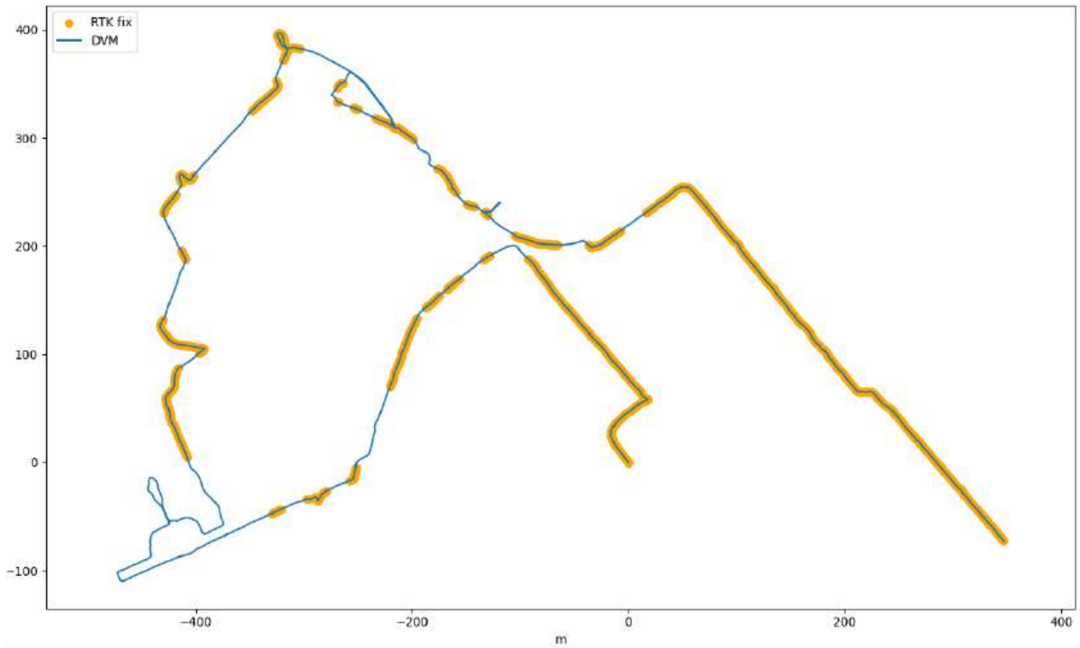


Figure 4: Trajectory of a difficult GNSS condition survey of 30 minutes. The distribution of PPK fixed of our proposed solution is shown in yellow. The start and end of the survey were done in good GNSS conditions but the biggest part of is done in harsh GNSS conditions.

Figure 5: Fix rate using GNSS-inertial tight coupling (Top), and using visual GNSS-tight coupling (bottom). Our proposed solution increases the fix rate by 4.5% over GNSS-inertial tight coupling. Only the fix rate in the difficult areas is shown. Our methods improve the fix quality on several points and it managed to fix the GNSS in the middle of long GNSS-fixed outage. These fixes in middle of outage are strategic and allow a centimetric accuracy in areas where it was not possible. Still there remain a big gaps with no fix that is easily associated to the area of the bottom left corner of the trajectory in Figure 4. The failure of getting a fix in the middle of this gaps can be caused by the too difficult GNSS condition or the quality of visual SLAM over this long 5min gaps being not precise enough for GNSS fix. The next section will assess the recovery precision of such long gaps.

b). Long GNSS Outage

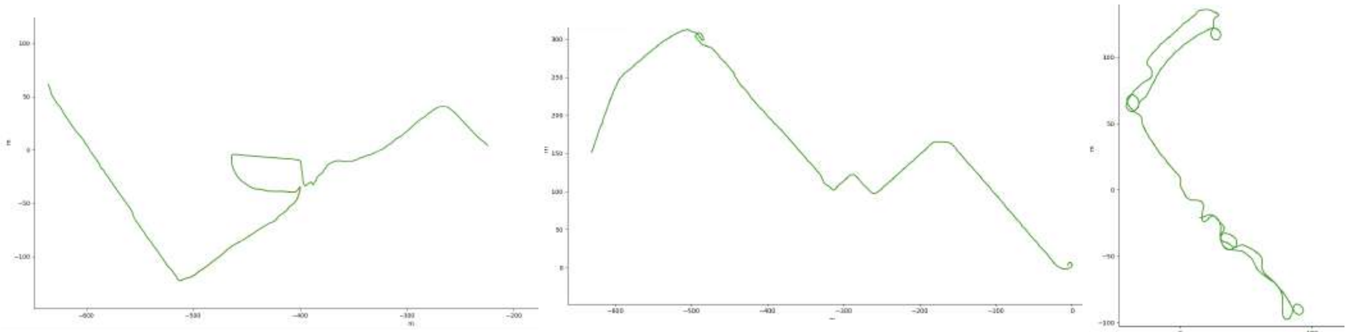


Figure 6: Trajectories of the GNSS outages tests. From left to right: residential test, in town test, parking test.

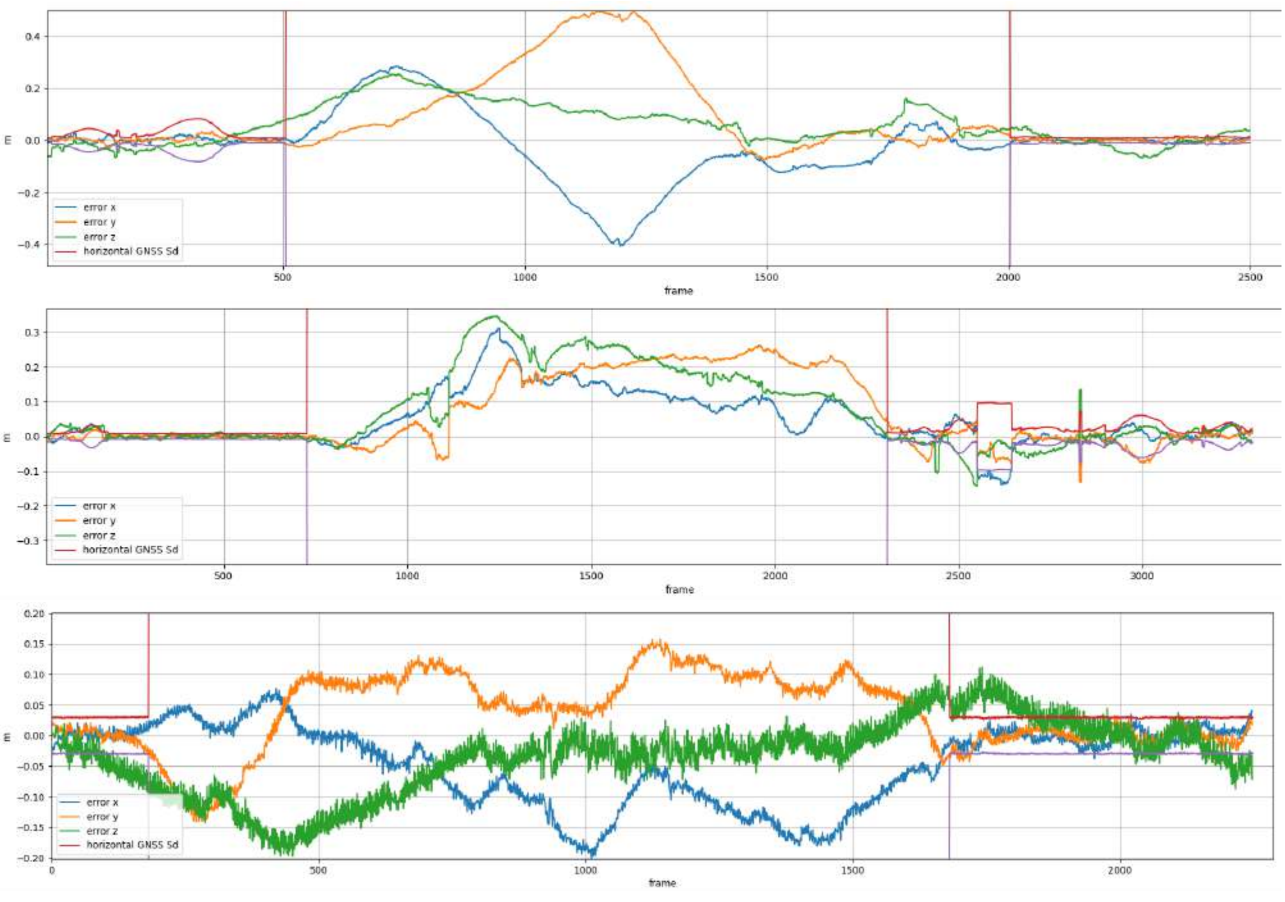


Figure 7: GNSS gaps simulation, From top to bottom: Residential dataset, in town dataset, parking dataset. The GNSS gaps are represented by the horizontal GNSS standard deviation in red and purple on the graphics.

This test will assess the capability of indoor positioning, where the GNSS is completely loss during a significant time, a time necessary to carry out an indoor survey for example.

We simulated on three datasets 5 minutes GNSS gaps, or a 450m long GNSS gap.

- The first test was done in a residential area with some narrow passages, and no loop closing. We had the lowest performance compared to the other tests with 50 cm maximum error (see Figure 7). This corresponds to 0.11 % of the traveled distance.

- The second test was done in town with no loop closing. The maximum error reached 35 cm (see Figure 7). This corresponds to 0.08 % of the traveled distance.
- The last test was done in the parking of SBG systems with some loop closing available. Vision seemed to take advantage of loop closing as the maximum error dropped to 20 cm (see Figure 7). This corresponds to 0.05 % of the traveled distance.

IV. LIDAR MAPPING

1. Acquisition Platform

The primary goal of our experimental setup was to compare several sensors for GNSS aided visual inertial navigation. We wanted a compact design that can be either embedded by a pedestrian, or carried by a vehicle. To demonstrate a practical application of such a solution, we embed a payload: A Velodyne LIDAR VLP-16. Hence our backpack is designed for simultaneous acquisition of several cameras and different IMUs. We embed two SBG IMUs (see table 1) and tested several cameras (see Table 2).

Table 1: IMUs.

INS	Picture	Grade	GNSS Receiver	Connectivity
Ellipse-D		7°/h bias instability	2 ublox zed-f9p	1 Serial port / 1 trigger
Quanta Micro Evaluation Kit		0.8°/h bias instability	2 ublox zed-f9p	3 serial port / 2 triggers / ethernet port / PTP / CAN

Table 2: Cameras.

Camera	Picture	Resolution	Characteristics	Connectivity
Intel D455		1.0 Mp x 2 color global shutter	90° FOV(H)	USB3
Intel T265		0.68 Mp x 2 gray global shutter	170° FOV(H)	USB3
Mynteye S		0.36 Mp x 2 gray global shutter	110° FOV(H)	USB3
ZED 2		2.2 Mp x 2 color rolling shutter	110° FOV(H)	USB3
Lucid Triton camera		5.0 Mp color global shutter	@ 110° FOV(H)	Gigabit Ethernet
E-con see3cam 20cug		2.0 Mp gray global shutter	@ 150° FOV(H)	USB3

Most cameras exhibit severe limitations:

- Some cameras emit lots of Electro Magnetic Radiations, jeopardizing GNSS. Schielding becomes indispensable.

- Rolling shutter cameras cannot be approximated by global shutter cameras. Whereas rolling shutter cameras can be approximated as global shutter for UAV photogrammetry. In case of pedestrian SLAM a dedicated algorithm for rolling shutter cameras seems mandatory.
- Effective resolution of color cameras is cut by 4 due to demosaicing. Redhibitory for small resolutions.
- Some cameras reached end of life and are no longer produced.

We finally decided to concentrate our tests on the most accurate setup we had: Quanta Micro with see3cam 20CUG. This setup is also of great research interest compared to available public dataset in the literature in term of ground truth availability, image resolution and IMU grade. We have a survey grade IMU, which provides accurate attitude to output quality LIDAR points cloud. The performance of the selected IMU is better but still comparable to the one available in the public dataset TUM-VI [17]. The camera selected has the biggest resolution among the Omnivision automotive grade monochrome global-shutter sensor. Which makes it cheaper and probably more precise than the industrial IDS uEye camera used in TUM-VI [17], since the resolution is bigger. Finally the ground truth is available on all open sky pieces of the trajectory. This allows us to simulate GNSS gaps and check angular and scale consistency on the long term. Such error cannot be checked on a limited ground truth such as the TUM-VI [17] one.

Table 3: Comparison of our backpack dataset with the hand-held TUM-VI dataset [17].

Dataset	Picture	Gyro Bias Instability	Camera	Ground Truth	Length per Log
TUM VI [17]		8°/h	1.0 Mp x 2 gray global shutter 150° FOV	1mm in static only at the beginning and the end	up to 30 min
Ours		0.8°/h	2.0 Mp x 2 gray global shutter 150° FOV	1cm in dynamic on the full open sky trajectory	up to 1 hour

We use a plastic suitcase for our setup as shows the picture in table 3, it has several advantages:

- It brings the rigidity to the sensor set.
- It is very modular, can easily be adapted to new sensors.
- The suitcase can either be carried as a backpack or attached to a vehicle.
- It protects the sensors inside and can be made water proof.

Our backpack can embed various batteries adapted to the current sensors in use. It embeds a laptop for real time GNSS aided visual inertial SLAM and for logging the data coming from all the different sensors.

2. LIDAR Calibration

Yet an additional work needs to be done to exploit our payload. The extrinsic calibration between the LIDAR and the INS is needed.

a). Related Work

Lidar-INS calibration has been well studied in the literature however manual procedures continue to be used. In a backpack context, [1] measures the lever arm from a blueprint and estimates the boresight angle using a comparison between INS gravity direction and floor normal direction with manual selection of ground floor. Several automatic methods for INS Lidar calibration use point to plane optimization. [18] automatically selects adaptive size planar patches using a quatree. [19] detects roofs in UAV lidar scans and use these roofs for point to plane optimization. Iterated Closest Point on plane also showed good performance in lidar SLAM context [20] [21]. This method is very effective in the general lidar context.

Parametric modeling of the planes to fit was used in [22], greatly increasing the complexity of the algorithm. [22] still experienced problems when trying calibration in a soccer-field, where the geometrical information is low. They also failed to calibrate an hypothetical lidar range offset. Hence, more care should be taken on other aspect of the algorithm such as plane classification, point association and outlier rejection.

b). Proposed Algorithm

Lidar calibration has been mostly explored in UAV context. We implement such a point to plane based boresight and lever arm estimation in a backpack context. We solve a non linear least square that optimizes the lever arm and boresight angle in order to minimize the distance of points to a plane. Instead of plane parametrization, we take into account the average plane displacement that depends of the optimization parameters (boresight and lever arm). Hence our non linear least square is:

$$dx = \left(\sum_k H_k^T W_k^{-1} H_k \right)^{-1} \left(\sum_k H_k^T W_k^{-1} h_k(x) \right) \tag{4}$$

With $h_k = f_k(x) - \overline{f_{i \in \mathbb{P}}(x)}$ and $H_k = \frac{\partial f_k}{\partial x}(x) - \overline{\frac{\partial f_{i \in \mathbb{P}}}{\partial x}(x)}$ and $f_k(x)$ the function that transform the point number k to the georeferenced frame using the boresight and lever arm as parameter. Such function can be found in [23]. This function is projected on the normal of the plane the point belongs to.

Our plane detection and point to plane association is very effective, hence we can reassociate points and planes at each Gauss Newton iteration. These steps are repeated until convergence:

- Store georeferenced points in a grid of about one meter square.
- Compute barycenter and SVD of the covariance of the point cluster in each square meters in order to determine if a cluster is a plane and where is the plane.
- Compute a least-square iteration with dynamic point rejections based on quadratic loss function [24].

It usually converges within 5 iterations. During the convergence, one can appreciate the number of clusters classified as plane increasing, and the thickness of the planes decreasing as shown in figure 8.

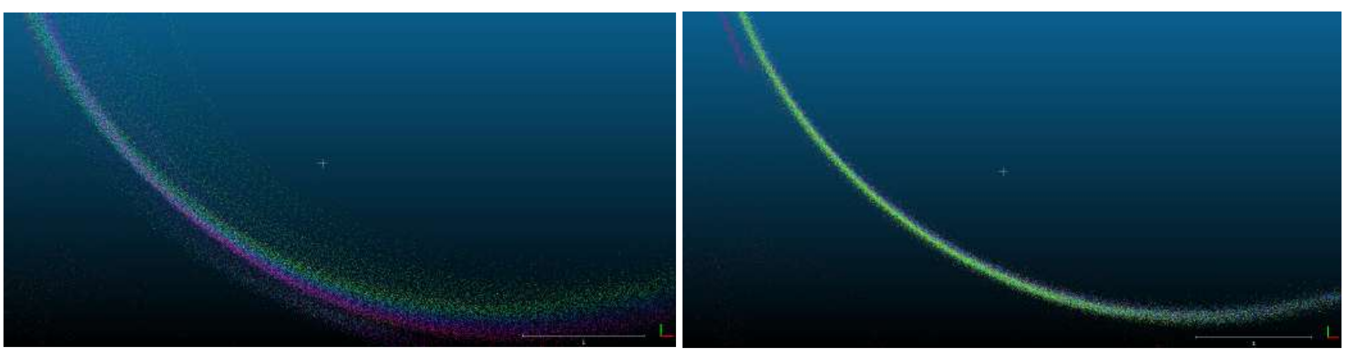


Figure 8: Point cloud of a water tank before calibration (left) and after calibration (right). Units in meters. The calibration parameters converged within 3 iterations. At the first iteration 2817 planes were used and at the last iteration 4965 plane were used. The boresight angle was corrected by 3° and the lever arm by 20cm. The point cloud used for calibration is plotted at the beginning of the paper.

V. DISCUSSION

When we look at the drift rate of our visual inertial SLAM, we have 0.1% to 0.05% of drift, this is much lower that the drift rate there are on the top ranking algorithms of the KITTI automotive visual odometry dataset [6] (1% to 0.5%). We can think that this can be due to to the automotive motion profile, introducing a high optical flow and blur. One can think of the absence of IMU data to help vision. However, from our experience running DVM on KITTI, the residuals are low, indicating an absence of blur and the motion profile of the automotive alleviates the need of an aiding sensor. Running the pedestrian visual-inertial dataset TUM-VI turned out to be more demanding. Together with the recent discovery of stereo rig rigidity issues and calibration issues [5] on the KITTI dataset, we can hope to reach a much lower drift rate with an automotive setup that does not suffer from mechanical issues. This would broaden the applications of our proposed solution.

VI. CONCLUSION

This paper presents a tight visual inertial GNSS processing able to provide a precise enough trajectory for mapping applications. We demonstrated that a single camera can do the work of Lidar when it comes to Simultaneous Localization and Mapping. We also showed that a tight coupling approach can help to recover a global centimetric trajectory thanks to the recovery of GNSS fixes in strategic areas of the trajectory. We point out that a camera typically used for object detection or collision detection is most of the time not suitable for visual SLAM. However, we find reliable, high-quality, and inexpensive cameras among automotive grade sensors that are very suited to visual SLAM. This makes our proposed solution very appealing.

REFERENCES

[1] M. Velas, M. Spanel, T. Sleziak, J. Habrovec, and A. Herout, "Indoor and outdoor backpack mapping with calibrated pair of velodyne lidars," *Sensors*, vol. 19, no. 18, p. 3944, 2019.

[2] M. Saidani, P. Sarri, A. Guinamard, and D. G. Maya, "Virtual base station algorithm and performance assessment," in *Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020)*, 2020, pp. 2696–2709.

[3] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. Montiel, and J. D. Tardós, "Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam," *IEEE Transactions on Robotics*, 2021.

[4] P. Bénet and A. Guinamard, "Deterministic visual mapping," in *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, 2021, pp. 3173–3188.

[5] I. Cvišić, I. Marković, and I. Petrović, "Soft2: Stereo visual odometry for road vehicles based on a point-to-epipolar-line metric," *IEEE Transactions on Robotics*, 2022.

[6] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *2012 IEEE conference on computer vision and pattern recognition*. IEEE, 2012, pp. 3354–3361.

[7] C. S. Fraser, "Automatic camera calibration in close range photogrammetry," *Photogrammetric Engineering & Remote Sensing*, vol. 79, no. 4, pp. 381–388, 2013.

[8] S. Cao, X. Lu, and S. Shen, "Gvins: Tightly coupled gnss–visual–inertial fusion for smooth and consistent state estimation," *IEEE Transactions on Robotics*, 2022.

[9] T. Qin, J. Pan, S. Cao, and S. Shen, "A general optimization-based framework for local odometry estimation with multiple sensors," *arXiv preprint arXiv:1901.03638*, 2019.

[10] J. E. Yoder, P. A. Iannucci, L. Narula, and T. E. Humphreys, "Multi-antenna vision-and-inertial-aided cdgnss for micro aerial vehicle pose estimation," in *Radionavigation Laboratory Conference Proceedings*, 2020.

[11] D. P. Shepard and T. E. Humphreys, "High-precision globally-referenced position and attitude via a fusion of visual slam, carrier-phase-based gps, and inertial measurements," in *2014 IEEE/ION Position, Location and Navigation Symposium-PLANS 2014*. IEEE, 2014, pp. 1309–1328.

[12] T. Li, H. Zhang, Z. Gao, X. Niu, and N. El-Sheimy, "Tight fusion of a monocular camera, mems-imu, and single-frequency multi-gnss rtk for precise navigation in gnss-challenged environments," *Remote Sensing*, vol. 11, no. 6, p. 610, 2019.

[13] L. Von Stumberg, V. Usenko, and D. Cremers, "Direct sparse visual-inertial odometry using dynamic marginalization," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2510–2517.

[14] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[15] P. Bénet, F. Novella, M. Ponchart, P. Bosser, and B. Clement, "State-of-the-art of standalone accurate auv positioning-application to high resolution bathymetric surveys," in *OCEANS 2019-Marseille*. IEEE, 2019, pp. 1–10.

[16] M. C. Deans, "Maximally informative statistics for localization and mapping," in *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*, vol. 2. IEEE, 2002, pp. 1824–1829.

[17] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stückler, and D. Cremers, "The tum vi benchmark for evaluating visual-inertial odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1680–1687.

[18] R. Keyetieu and N. Seube, "Automatic data selection and boresight adjustment of lidar systems," *Remote Sensing*, vol. 11, no. 9, p. 1087, 2019.

[19] J. Skaloud and P. Schaer, “Towards automated lidar boresight self-calibration,” Tech. Rep., 2007.

[20] J. Zhang and S. Singh, “Loam: Lidar odometry and mapping in real-time.” in *Robotics: Science and Systems*, vol. 2, no. 9. Berkeley, CA, 2014, pp. 1–9.

[21] P. Dellenbach, J.-E. Deschaud, B. Jacquet, and F. Goulette, “Ct-icp: Real-time elastic lidar odometry with loop closure,” in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 5580–5586.

[22] J. Skaloud and D. Lichti, “Rigorous approach to bore-sight self-calibration in airborne laser scanning,” *ISPRS journal of photogrammetry and remote sensing*, vol. 61, no. 1, pp. 47–59, 2006.

[23] Z. Li, J. Tan, and H. Liu, “Rigorous boresight self-calibration of mobile and uav lidar scanning systems by strip adjustment,” *Remote Sensing*, vol. 11, no. 4, p. 442, 2019.

[24] J. Zubizarreta, I. Aguinaga, and J. M. M. Montiel, “Direct sparse mapping,” *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1363–1370, 2020.