

A GNSS-FREE UNMANNED AERIAL LASER SCANNING SYSTEM

M. Shahbazi^{1,*}, C. Cortes², P. Ménard¹, J.S. Bilodeau¹

¹ Centre de géomatique du Québec, Saguenay, G7H 1Z6 Canada – (mshahbazi, pmenard, jsbilodeau)@cgq.qc.ca

² Dept. of Geomatics Engineering, University of Calgary, University Drive NW, Calgary, T2N 1N4 Canada - cecortes@ucalgary.ca

Technical Commission II

KEY WORDS: LiDAR, Sensor Fusion, System Integration, Platform Calibration, Georeferencing

ABSTRACT:

In this paper, the procedure of developing and evaluating a UAV-borne mapping system is described. The system is equipped with both a LiDAR and a camera. The system mounting parameters, as well as the intrinsic parameters of the individual sensors, are calibrated rigorously. Simultaneous calibration of the LiDAR intrinsic parameters and the LiDAR-camera mounting parameters is performed in a self-calibrating bundle adjustment with additional relative orientation constraints. A visual-inertial approach is proposed to georeference the laser scans without using a GNSS receiver. This approach is motivated not only by the interest of users in low-cost systems but also by the fact that the integrity of GNSS signals might be affected under several environmental conditions, e.g., indoors, in urban canyons, under tree canopies. It is shown that a low-cost inertial measurement unit not equipped with a dual-frequency, real-time kinematic GNSS receiver is still useful for georeferencing the laser scanning data with cm-level accuracy. The scans are also textured using the images captured by the camera, which enriches the LiDAR point clouds with spectral information.

1. INTRODUCTION

Over the past few years, the interest in low-altitude aerial mapping using camera-equipped Unmanned Aerial Vehicles (UAVs) has exponentially grown (Coops et al., 2019; Kerle et al., 2019; Ren et al., 2019; Shahbazi et al., 2014; Yao et al., 2019; Mogili and Deepak, 2018). This rapid growth is mainly due to the significant advances in aeronautics technologies, autonomous navigation methods, digital cameras, as well as photogrammetry and computer vision disciplines. Along with this trend, manufacturers of three-dimensional (3D) laser scanners have also changed their focus to the UAV industry. Consequently, affordable LiDAR sensors designed for integration with drone platforms are available nowadays. Multi-sensor systems equipped with both camera and LiDAR offer the advantages of both photogrammetry and ranging techniques (Nouwakpo et al., 2016; Wallace et al., 2012). As a result, UAVs equipped with both a camera and a LiDAR sensor are ideal to cover a large variety of applications in different environmental conditions, e.g., forestry, wildlife research, structural health monitoring, precision agriculture, and corridor mapping.

This is the primary motivation behind developing a multi-sensor solution (Figure 1) in this study. The proposed system includes an optical digital camera (a6000, Sony), a 3D lidar (VLP16, Velodyne LiDAR Inc), a low-cost inertial navigation system (INS) (VN200, VectorNav Technologies), a compact processor, electronic components, and the casing. The total cost of building this integrated system is under 15K Canadian dollars, which makes it an affordable solution for small and medium enterprises. The proposed solution involves the rigorous calibration of the system, modeling intrinsic calibration parameters of the sensors, geo-referencing the laser scanning data, and fusing data from passive and active sensors. Considerable effort is put into 1) the geometric calibration of the system; and 2) developing an error-state Kalman filter for fusing the visual information obtained through photogrammetry

with the raw measurements of the inertial measurement unit (IMU) while simultaneously estimating the mounting parameters between the IMU and the camera.

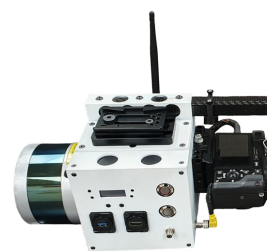


Figure 1. The in-house built multi-sensor system

2. CALIBRATION

In the case of a LiDAR-camera system, three basic types of geometric calibration are required: intrinsic camera calibration, intrinsic LiDAR calibration, extrinsic system calibration. In this paper, several calibration approaches are investigated and compared: **(A)** Individual camera calibration in a free-network bundle adjustment (BA), individual LiDAR intrinsic calibration in a controlled BA, controlled BA for estimating the exterior parameters (EOPs) of the images taken together with the LiDAR, separate calculation of the mounting parameters (the bore-sight angles and the lever-arm) between the two sensors directly from their estimated EOPs; **(B)** Individual camera calibration and individual LiDAR intrinsic calibration in separate free-network BAs, calculation of the mounting parameters in a controlled BA with additional relative-orientation (RO) constraints; **(C)** Individual camera calibration in a free-network BA, integrated controlled self-calibrating BA with RO-constraints for simultaneously estimating both the system mounting parameters and the intrinsic calibration parameters of the LiDAR. All the bundle adjustments are solved in a sparse fashion, following our previous approach (Shahbazi et al., 2017), in order to handle a large number of observations.

* Corresponding author

Intrinsic camera calibration involves modeling the lens radial distortions, lens decentering distortions, and interior orientation parameters (principal point offsets and focal length). Equation (1) describes the augmented collinearity equations used in the bundle adjustments to model the calibration parameters of the camera. In addition to calibrating the system's camera (Sony camera) an auxiliary camera (a Canon EOS M5 camera with a 24-mm lens) was also calibrated. This additional camera was used to extract the approximate mounting parameters between the camera and the IMU using the approach described in our previous work (Cortes et al., 2018).

$$\begin{pmatrix} x_{i,j} - x_p \\ y_{i,j} - y_p \end{pmatrix} + f \begin{pmatrix} x'_{i,j} (1 + K_1 r_{i,j}^2 + K_2 r_{i,j}^4 + K_3 r_{i,j}^6) + \\ 2P_1 x'_{i,j} y'_{i,j} + P_2 (r_{i,j}^2 + 2x_{i,j}^2) \\ y'_{i,j} (1 + K_1 r_{i,j}^2 + K_2 r_{i,j}^4 + K_3 r_{i,j}^6) + \\ 2P_2 x'_{i,j} y'_{i,j} + P_1 (r_{i,j}^2 + 2y_{i,j}^2) \end{pmatrix} = 0 \quad (1)$$

where

$$\begin{bmatrix} u_{i,j} & v_{i,j} & w_{i,j} \end{bmatrix}^T = \begin{bmatrix} \mathbf{R}_e^{c_j} & -\mathbf{R}_e^{c_j} \mathbf{r}_{c_j}^e \end{bmatrix} \tilde{\mathbf{X}}_i^e$$

$$x'_{i,j} = \frac{u_{i,j}}{w_{i,j}}, \quad y'_{i,j} = \frac{v_{i,j}}{w_{i,j}}$$

$$r_{i,j}^2 = x_{i,j}^2 + y_{i,j}^2$$

k_1, k_2, k_3 : radial lens distortion coefficients
 p_1, p_2 : decentering lens distortion coefficients
 i : object point index
 j : image index
 e : denoting object/ground coordinate system

In Equation (1), $\mathbf{R}_e^{c_j}(\omega_j, \phi_j, \kappa_j)$ is the 3D rotation matrix from the object to the image coordinate system, where $\omega_j, \phi_j, \kappa_j$ represent the rotation angles around the z, y and x axes, respectively. $\mathbf{r}_{c_j}^e$ is the 3D position of the perspective centre resolved in the object coordinate system. $\tilde{\mathbf{X}}_i^e$ is the homogeneous coordinates of the object point. Coordinates $(x_{i,j}, y_{i,j})$ are the observations of the point in the image principal coordinate system. Interior orientation parameters of the camera include (x_p, y_p, f) .

To perform the calibrations, a multi-façade, multi-resolution test-field with appropriate types of targets was used (Figure 2). The test-field includes 19 planar objects, a checkerboard pattern with 72 corners, and a total of 111 circular targets. The planar features in the LiDAR scan are detected using region-growing segmentation with RanSAC. Identifying corresponding planar features that are observed by the LiDAR from different viewpoints is completed by back projecting the co-planar points into the image space followed by segment matching in the image space.

Intrinsic LiDAR calibration involves modeling the systematic effects of rangefinders' offsets, horizontal angle offsets, and vertical circle zero index errors for each of the laser beams in the VLP16. In total, 16 rangefinder offsets, 15 horizontal angle offsets and 15 vertical zero index error terms are considered as

intrinsic calibration parameters since the horizontal angle offset and the vertical circle zero index error term of one laser need to be constant to define the scanner space (Chan, 2015). The observation equations used in the BA for LiDAR measurements are as follows, where $\mathbf{R}_e^{s_j}, \mathbf{r}_{s_j}^e$ represent the EOPs of the LiDAR.

$$\begin{aligned} u_{i,j} - (\rho_{i,j} + \delta_{\rho_L}) \cos(\theta_L + \delta_{\theta_L}) \sin(\alpha_{i,j} + \delta_{\alpha_L}) &= 0 \\ v_{i,j} - (\rho_{i,j} + \delta_{\rho_L}) \cos(\theta_L + \delta_{\theta_L}) \cos(\alpha_{i,j} + \delta_{\alpha_L}) &= 0 \\ w_{i,j} (\rho_{i,j} + \delta_{\rho_L}) \sin(\theta_L + \delta_{\theta_L}) &= 0 \end{aligned} \quad (2)$$

where:

$$\begin{bmatrix} u_{i,j} & v_{i,j} & w_{i,j} \end{bmatrix}^T = \begin{bmatrix} \mathbf{R}_e^{s_j} & -\mathbf{R}_e^{s_j} \mathbf{r}_{s_j}^e \end{bmatrix} \tilde{\mathbf{X}}_i^e$$

ρ : range
 α : horizontal angle
 θ_L : vertical angle
 L : laser id $\in [0, 15]$
 δ_{ρ_L} : range zero-error
 δ_{α_L} : horizontal circle error
 δ_{θ_L} : vertical circle zero index error

For calibration approaches B and C, RO stability constraints are added to the BA. At a station, k , the relative orientations $\mathbf{R}_s^c, \mathbf{r}_s^c$ between the camera and the LiDAR can be denoted as follows.

$$\begin{aligned} \mathbf{R}_s^c &= \mathbf{R}_e^{c_k} (\mathbf{R}_e^{s_k})^T \\ \mathbf{r}_s^c &= \mathbf{R}_e^{c_k} (\mathbf{r}_{s_k}^e - \mathbf{r}_{c_k}^e) \end{aligned} \quad (3)$$

This equation should hold for all the stations. That is, at a different station, h , the stability of the RO parameters obliges Equation (4). Despite being in the form of translation vectors, these equations involve rotation matrices. Thus, they constrain both the lever arm offsets and the boresight angles.

$$\begin{aligned} \mathbf{R}_e^{s_h} (\mathbf{r}_{c_h}^e - \mathbf{r}_{s_h}^e) &= \mathbf{R}_e^{s_k} (\mathbf{r}_{c_k}^e - \mathbf{r}_{s_k}^e) \\ \mathbf{R}_e^{c_h} (\mathbf{r}_{s_h}^e - \mathbf{r}_{c_h}^e) &= \mathbf{R}_e^{c_k} (\mathbf{r}_{s_k}^e - \mathbf{r}_{c_k}^e) \end{aligned} \quad (4)$$

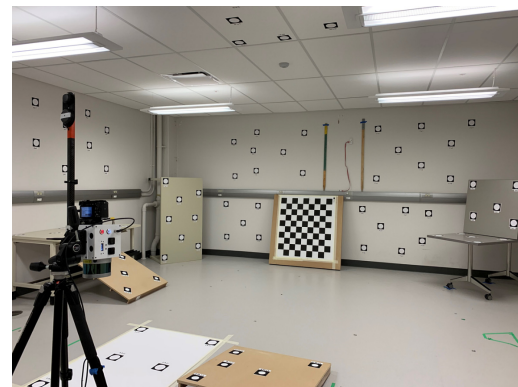


Figure 2. Front view of the calibration test-field

3. GEO-REFERENCING

The ultimate objective of georeferencing is to produce dense, textured point clouds from LiDAR scans in a mapping reference coordinate system. In our system, direct georeferencing using

GNSS/INS is not possible due to the low accuracy of the low-cost sensors (2-5m in position and 0.1-0.5 degrees in attitude). As such, GNSS position and velocity observations are ignored. Instead, the camera is used as a pose estimator. That is, the camera pose is loosely coupled with the inertial measurements of the IMU, to not only compensate inertial sensors' biases but also to estimate the uncompensated errors in the mounting parameters between the camera and the IMU. The outcome of this sensor fusion is then used to georeference the LiDAR scans and assign RGB color values to the point cloud.

The geometric relations between the sensors are represented in Figure 3. In Figure 3, the symbols "c", "s", "b", and "e" denote various frames involved in the mapping process:

- c : image principal coordinate system (c-frame)
- s : sensor coordinate system of LiDAR (s-frame)
- b : INS body-frame (b-frame)
- e : object coordinate system

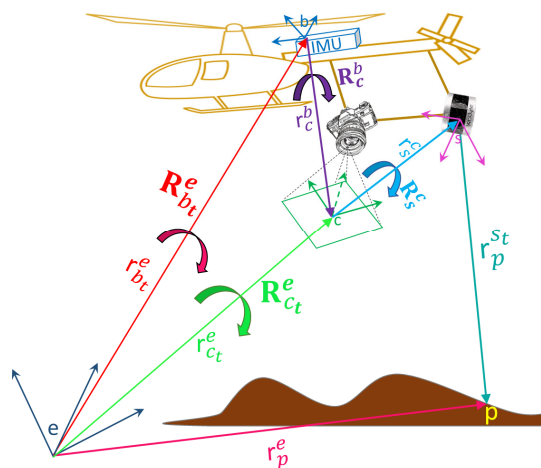


Figure 3. Geometric relations between the sensors and their measurements

The translation vectors and the rotation matrices of Figure 3 are defined below.

- $r_{c_t}^e$: coordinates of the perspective center in the object coord. system at time "t"
- r_c^b : lever-arm offset between camer and IMU
- r_s^c : lever-arm offset between LiDAR and camera
- $r_{b_t}^e$: IMU position in the object coord. system at time "t"
- $r_p^{s_t}$: LiDAR scan of an object point at time "t"
- r_p^e : absolute 3D coordinates of the object point
- $R_{c_t}^e$: rotation from the camera to the object coord. system at time "t"
- $R_{b_t}^e$: rotation from IMU body-frame to the object coord. system at time "t"
- R_c^b : boresight matrix between the camera and IMU
- R_s^c : boresight matrix between the LiDAR and camera

The georeferenced position of a point measured by the laser scanner at time t can be derived via Equation (5).

$$r_p^e = r_{b_t}^e + R_{b_t}^e r_c^b + R_{b_t}^e R_c^b r_s^c + R_{b_t}^e R_c^b R_s^c r_p^{s_t} \quad (5)$$

In Equation (5), $R_{b_t}^e, r_{b_t}^e$ are the outcomes of the sensor fusion technique that will be explained later. R_c^b, r_c^b are approximately estimated through an offline calibration and are refined through the proposed sensor fusion technique. R_s^c, r_s^c are estimated through the offline calibration. And, $r_p^{s_t}$ contains the intrinsically calibrated measurements of the LiDAR.

The first step to this methodology is determining the EOPs of the images in the same coordinate system, in which the LiDAR point cloud needs to be created. Here, indirect geo-referencing is performed. An adequate number of ground control points (GCPs) are established in the scene. Their coordinates are accurately measured in the mapping coordinate system. The following photogrammetric processing workflow (Figure 4) is applied to the images using commercial software, Pix4D Mapper Pro. The inputs to this procedure include the IOPs and intrinsic calibration parameters of the camera, as well as the detected and labelled GCPs in all images. Detection and labelling of the GCPs is performed using an automatic approach suggested by Shahbazi et al. (2015). The outputs of interest from this procedure are the EOPs of the images (R_c^e, r_c^e).

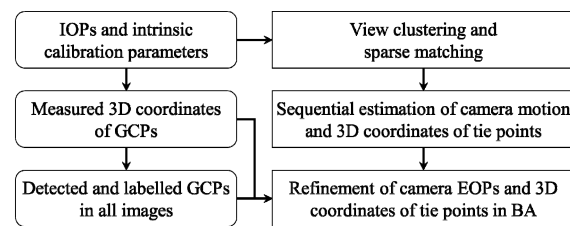


Figure 4. Photogrammetric workflow to estimate the EOPs of images

3.1 Camera-IMU fusion

The approach used in this thesis is based on a Kalman Filter (KF) for fusing the IMU data with the visual inputs (EOPs of images) (Lynen et al., 2013). In this approach, the camera is used as an additional sensor to compensate for the temporal drifts of the IMU. In other words, the IMU is the process/propagation sensor, and the camera is the update sensor. In this approach, while it is assumed that the intrinsic camera parameters are known and fixed (Section 2), the mounting parameters describing the lever arm offsets and boresight angles between the IMU and the camera are only known approximately and need to be re-estimated. Studies have proven that these mounting parameters are observable regardless of the linear motion of the IMU-camera rig (Mirzaei and Roumeliotis, 2008). There are two different ways to combine inertial and visual measurements for high-frequency pose estimation. The first one is the loosely coupled approach, which considers the inertial sensor and the vision system as two separate modules that only exchange information when possible (Weiss and Siegwart, 2011). The second approach is the tightly coupled approach, which combines the inertial and the vision units into a single filter (Kelly and Sukhatme, 2011). In this study, loosely coupled integration is applied. Decoupling the vision system from the inertial system allows us first to apply the full photogrammetric procedure separately, and then use the refined EOPs for integration with the inertial measurements. As such, visual scale or attitude drifts in the vision system can be resolved separately

and avoided from affecting the inertial system. Besides, in tightly coupled solutions, the computational complexity varies by the number of observed image features. In contrast, in the loosely coupled approach, the computational complexity is constant since the observations are always of the same dimension (6 degrees of freedom in EOPs). The state of the filter is composed of the position \mathbf{r}_b^n of the IMU; its velocity \mathbf{v}_b^n ; its attitude quaternion \mathbf{q}_b^n describing a rotation from b-frame to n-frame, where n-frame is the local navigation frame; gyro and acceleration biases ω_b and \mathbf{a}_b ; acceleration due to gravity \mathbf{g}^n that can capture the misalignments of the nominal n-frame horizontal plane with the true n-frame; this is especially useful when complete information of the mapping reference frame is unavailable (Lupton and Sukkarieh, 2009); boresight quaternion \mathbf{q}_c^b describing a rotation from the c-frame to the b-frame; and the position \mathbf{r}_c^b of the camera perspective center in the b-frame. The entire state results in a 26-element state vector as follows.

$$\mathbf{X} = \{\mathbf{r}_b^n, \mathbf{v}_b^n, \mathbf{q}_b^n, \omega_b, \mathbf{a}_b, \mathbf{g}^n, \mathbf{q}_c^b, \mathbf{r}_c^b\} \quad (6)$$

The nominal state of the system is, for the most part (position, velocity, attitude), integrable in a non-linear fashion. Therefore, their linearization can result in significant errors if a conventional extended Kalman filter is used. However, the error state includes small-signal quantities in terms of noises and perturbations. Therefore, it evolves through a linear dynamic system and is suitable for Kalman filtering. The applied error-state filter is summarized in Figure 5.

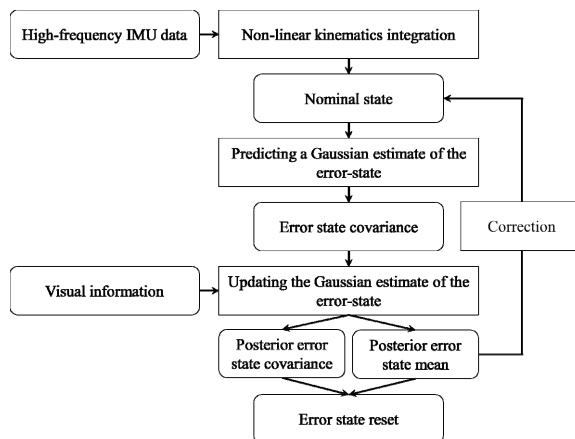


Figure 5. Error-state Kalman filtering for fusing IMU data with visual information obtained from photogrammetry

3.2 Point cloud texturing

Once the INS pose is estimated through the method of Section 3.1, Equation (5) can be used to georeference the LiDAR scans. To assign a color to each point, then the following algorithm is used. The inputs to this algorithm include the LiDAR point cloud, the images that are rectified for all non-linear distortions, the camera IOPs, and the EOPs of the images.

Algorithm 1: point cloud texturing

For every point \mathbf{P} in the point cloud
 Find the five closest images to $\mathbf{P} \rightarrow$ image set Δ
 For every image \mathbf{I} in set Δ

If \mathbf{P} is in front of \mathbf{I} and in the field of view of \mathbf{I}
 Back project \mathbf{P} to $\mathbf{I} \rightarrow \mathbf{p}$
 Use the nearest-neighbor method to interpolate a color value for \mathbf{p}
 Calculate the radial distance of \mathbf{p} from the image center $\rightarrow \mathbf{d}$
 Assign \mathbf{d} to \mathbf{I}

Else
 Remove \mathbf{I} from Δ

If Δ is not empty
 Select the image in Δ with the shortest \mathbf{d} to assign a color to \mathbf{P}

4. EXPERIMENTS AND RESULTS

To calibrate the cameras using the test-field of Figure 2, a total of 100 control images with Sony A6000 and another 100 control images with Canon EOS M5 were captured from various locations and orientations, as shown in Figure 6. A total of 12994 targets were observed in these 200 images, resulting in a degree of freedom of 24226. The detected and labelled targets were all manually verified to reduce the risk of gross errors in the self-calibration BA. The BA was also equipped with a Huber loss function to minimize the effect of any undetected errors in the observations.

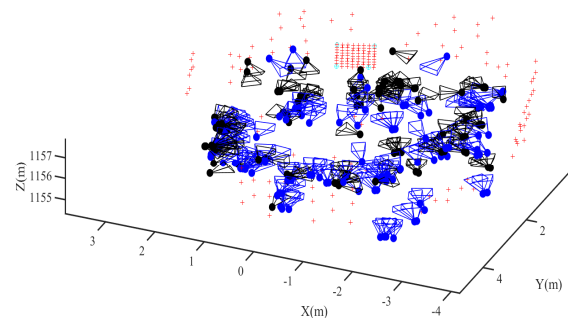


Figure 6. Imaging network configuration for intrinsic calibration of the camera; blue frames correspond to Sony images and black frames to Canon images

The resulted mean and standard deviation (StD) of the residuals on image observations after the self-calibration were $-0.2\text{e-}4$ and 0.10 pixels along the x-direction; $0.1\text{e-}4$ and 0.10 pixels along the y-direction; and, 0.12 and 0.08 pixels in magnitude. An independent set of 15 Canon images and 23 Sony images were collected for check purposes, as shown in Figure 7. A total of 3148 targets were observed in these 38 images. In the check calibrations, 4 extreme corners of the checkerboard were used as control points. The coordinates of all other checkerboard targets were used as checkpoints. In the check adjustment using the calibrated intrinsic parameters, the mean and StD of the residuals were $0.7\text{e-}4$ and 0.10 pixels along the x-direction; $-0.6\text{e-}4$ and 0.10 pixels along the y-direction; and, 0.12 and 0.08 pixels in magnitude. The mean and root mean square (RMS) of reprojection errors (calculated by back-projecting checkpoints to check images) along the x-direction are -0.01 and 0.02 pixels for Canon; and, 0.07 and 0.02 pixels for Sony images. The same statistics along the y-direction are -0.02 and 0.03 pixels for Canon images; and, 0.01 and 0.04 pixels for Sony images.

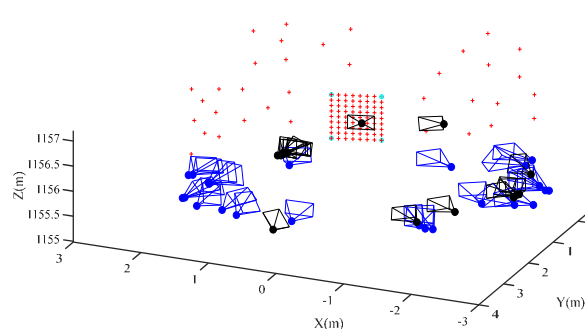


Figure 7. Configuration of check images

In order to calibrate the LiDAR intrinsically and also to estimate its mounting parameters with respect to the camera, LiDAR scans and camera images were captured from the test-field at 20 different stations (Figure 8). In Figure 8, the red dots are the scanned points by the LiDAR, and the cyan circles are the targets observed by the camera. A total of 164707 scanned planar points, and 922 target observations were made. That is, approximately 660672 observation equations were generated for the bundle adjustment. Therefore, sparse bundle adjustment became crucial. Check data was collected from a total of 10 independent stations to test the accuracy of the calibration procedure (Figure 9). In the period between the calibration and check, the system was also used outdoors for data collection.

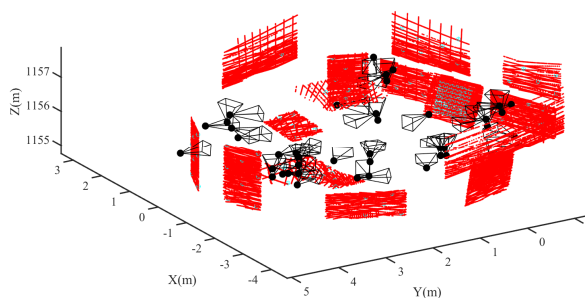


Figure 8. Data collection network for LiDAR calibrations

To be able to verify the accuracy of each calibration approach, ground-truth data of the planar features were captured using a FARO Focus 3D terrestrial laser scanner (Figure 10). As a phase-based laser scanner, it has superior accuracy to VLP16. Also, it can be set to acquire data with high density. To avoid any unnecessary errors in the ground-truth data due to multi-scan registrations, the FARO scanner was placed in a way that all the planes could be scanned from a single station.

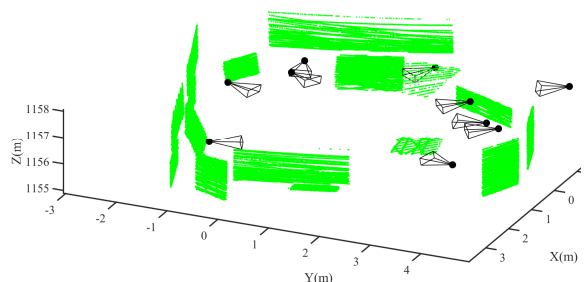


Figure 9. Data collection network for checking LiDAR calibrations

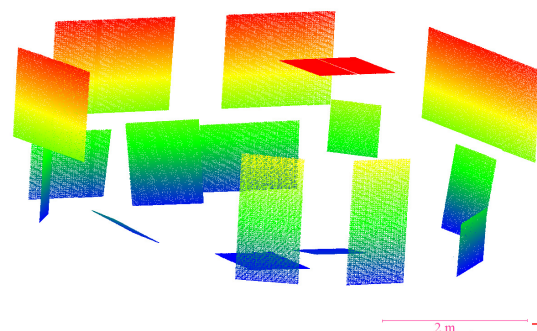


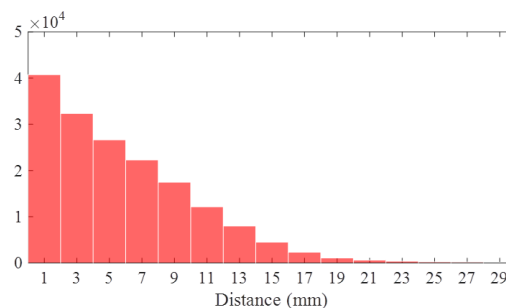
Figure 10. Ground-truth data collected by FARO laser scanner

The results from the three calibration approaches are summarized in Table 1. The standard deviation of the RO parameters between the LiDAR and the camera demonstrate the uncertainty in the calculated RO parameters, and the distance of the check point-cloud to the ground-truth point-cloud represents the accuracy of modeling.

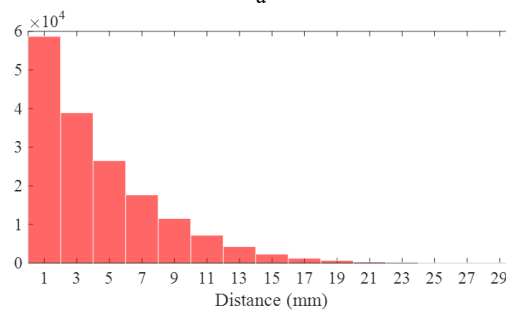
Calibration scenario	Lever-arm uncertainty (mm)	Boresight uncertainty (deg)	RMS 3D distance (mm)
A	6.50	0.1368	7.68
A- W/O intrinsic LiDAR calibration	7.74	0.1836	8.87
B	4.42	0.0700	6.28
B- W/O intrinsic LiDAR calibration	5.55	0.0925	7.32
C	3.99	0.0587	5.87

Table 1. LiDAR Calibration results

Figure 11 summarizes the distances between the check point-clouds and the ground-truth point clouds for all the calibration approaches in the form of histograms.



a



b

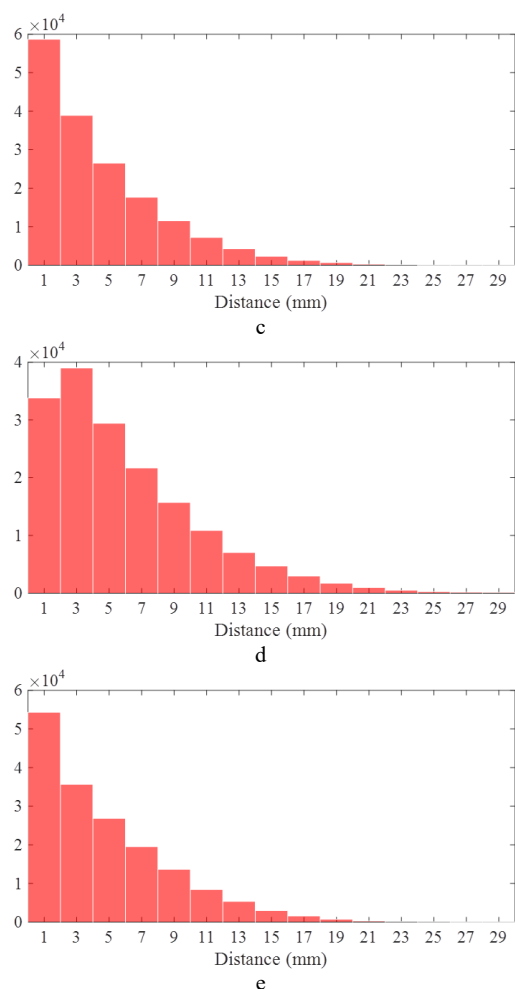


Figure 11. Histograms of the distance between the check point-cloud and the ground-truth point-cloud with a) calibration approach A, b) in approach B, c) in approach C, d) in approach A without intrinsic calibration of the LiDAR, and e) approach B without intrinsic calibration of LiDAR

In general, it can be seen that intrinsic calibration of the LiDAR sensor helps reducing the modeling errors by 14.2% (approach B with and without an intrinsic calibration) to 13.4% (approach A with and without an intrinsic calibration). Using a RO-constrained bundle adjustment reduces the uncertainty in the RO parameters by 48.8% (approach B compared to approach A). Calibrating the LiDAR parameters along with the RO constraints (approach C compared to approach B) allows reducing the modeling errors by 6.5% and the RO uncertainties by 16.1%. As a result, the calibration approach C yields the best results.

In order to test the georeferencing approach, the system was installed on a UAV (Figure 12), and the trajectory of Figure 13 was flown over an agricultural field. Images were captured with a frequency of 1.5 fps. The rotation rate of the LiDAR was set to 600 rotations per minute. Distance and azimuth thresholds were applied to avoid scanning points that were too far from the trajectory. The IMU logging frequency was set to 67 Hz, which was the maximum value possible considering the baud rate limitations of the embedded computer. Due to unpredicted snowfall and dense cloud covers on the day of data collection, the field was sparsely covered by snow patches, and images

were considerably underexposed. We could neither decrease the shutter speed (to avoid motion blur) nor increase the aperture opening (to prevent out-of-focus blur) to fix the exposure issue. Due to the sudden snowfall that day, thorough control data (e.g., terrestrial reference scans or abundant checkpoints) could not be acquired either.



Figure 12. Integration of the LiDAR-camera system onboard a UAV

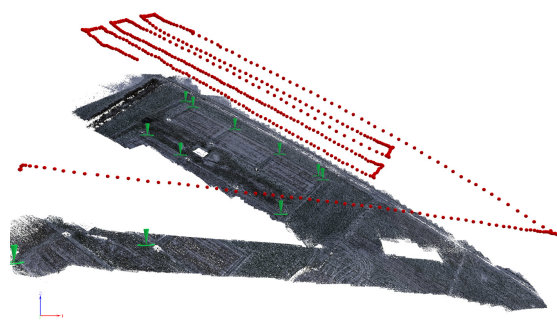


Figure 13. Position of captured images (red dots) and location of the ground control points (green pins) shown in the 3D view of the study area

Five of the ground points were used as GCPs, and the other six were used as checkpoints. GCPs coordinate were measured with real-time kinematic GNSS receiver. The estimated coordinates of the checkpoints differ from their measured ones by an RMS of 4.2 cm. We believe the snow, which fell in between the time of image acquisition and ground-truth collection, had caused errors in the measurements. This could explain a portion of the errors on the checkpoints.

Using the proposed method, camera-IMU fusion was performed to estimate an accurate trajectory. Figure 14 shows the trajectory determined by camera-aided INS (proposed method of Section 3.1) and GNSS-aided INS (Manufacturer's software). To measure the accuracy of the estimated trajectory, every 10th image in the capture sequence escaped the Kalman correction phase. The photogrammetrically measured EOPs of these images were reserved as check poses. Figure 15 shows the errors of the camera-aided INS and the GNSS-aided INS at the check poses. The denoted errors in position and orientation are the maximum of the errors in all three dimensions (xyz). The RMS errors of camera-aided INS are 1.59 cm in position and 0.0022° in orientation. The RMS errors of GNSS-aided INS are 190.07 cm and 19.07°.

In Figure 14, it should be noted that in camera-IMU fusions, the edges of the trajectory during the sharp turns were cut out to avoid the noisy trajectory from affecting the LiDAR georeferencing accuracy.

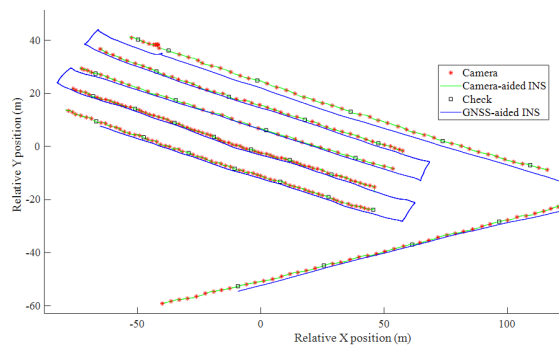


Figure 14. UAV trajectory: green color represents the camera-aided INS estimations

The LiDAR scans were georeferenced using this camera-aided, INS data and textured with the method of Section 3.2. Then, median filtering was applied to remove the noise in the final point cloud (Figure 16). Figure 17 shows the reflectivity map observed by the LiDAR. Since no ground-truth point cloud could be acquired from the study area, a dense point cloud was generated using photogrammetric processing of the images. The photogrammetric point cloud was used as a reference to measure the quality of the LiDAR point cloud.

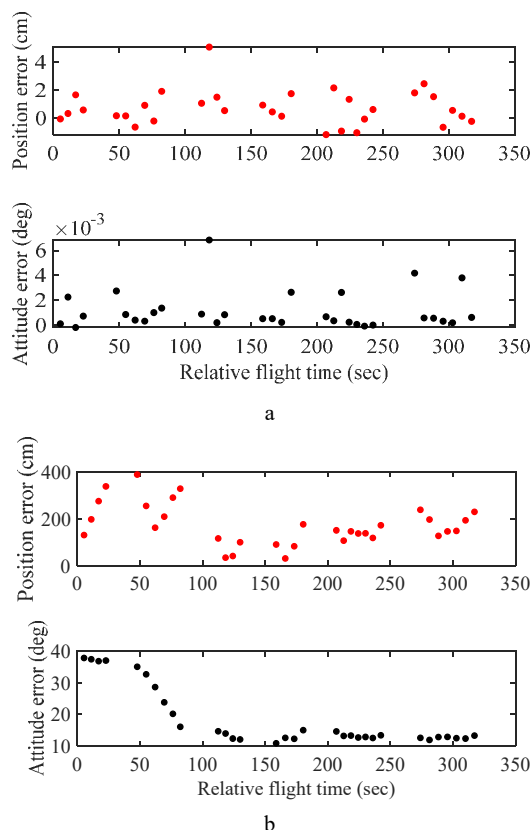


Figure 15. Pose-estimation errors by a) camera-aided INS and b) GNSS-aided INS

Comparing the LiDAR point cloud to the photogrammetric point cloud reveals that there is only a difference of few centimeters between the two point-clouds. The mean and RMS of these distances are 4.49 cm and 5.72 cm, respectively. Most distances between the two point-clouds are below 15 cm. We believe that substantial differences correspond mainly to features that the photogrammetric point cloud could not have

captured due to the combination of snow coverage and underexposure. However, these features were present in the LiDAR scans. The distance of the LiDAR point cloud from the photogrammetric point cloud is shown in Figure 19, and its histogram is summarized in Figure 18.



Figure 16. 3D view of the point cloud generated by LiDAR

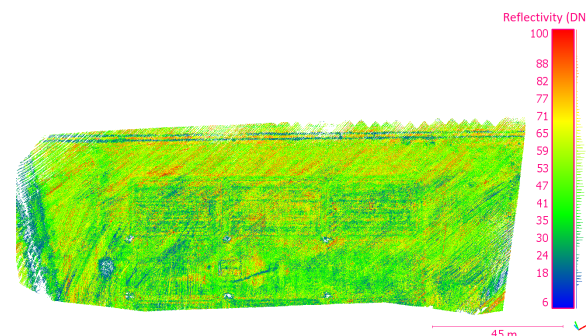


Figure 17. LiDAR point cloud with observed reflectivity values

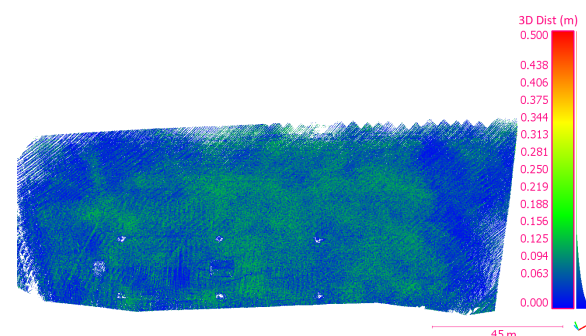


Figure 18. Absolute distance of the LiDAR point cloud from the photogrammetric point cloud

5. CONCLUSIONS

In this paper, a UAV-based mapping system equipped with a camera and a LiDAR sensor was described. It was shown that geometric calibration of the system mounting parameters and the intrinsic parameters of the individual sensors could improve the scanning accuracy considerably. A low-cost IMU was used for direct georeferencing of the laser scanning data. A visual-inertial, error-state Kalman filtering approach was proposed to compensate for both the biases of the inertial sensors and the mounting misalignments between the camera and the IMU. This approach eliminated the need for an expensive high-performance GNSS/INS system. Yet, high direct georeferencing accuracy could be achieved (better than 15 cm). Furthermore,

the scans were textured using the captured imagery, enriching the LiDAR point cloud with the spectral information of the images.

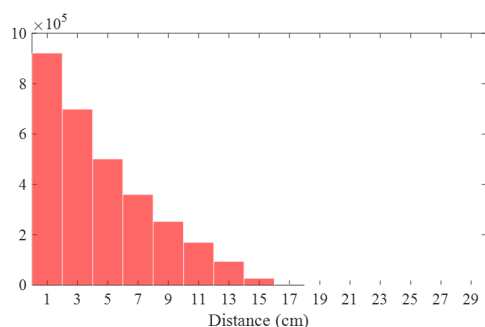


Figure 19. Histogram of the distances between the LiDAR point cloud and the photogrammetric point cloud

ACKNOWLEDGEMENTS

This research was supported by multiple funding resources including NSERC Discovery Grant, Mitacs Accelerate, Canada's Tri-Council Agencies New Frontiers in Research Fund, and NSERC College and Community Applied Research and Development Grant. The authors would like to thank Dr. Karem Chokmani from L'Institut national de la recherche scientifique (INRS) for both financially supporting the purchase of the equipment used in this study and for providing their laboratory drone in the tests. The authors would also like to acknowledge the assistance received from the undergraduate interns, Eric Ho and Erica Lemieux from the University of Calgary, during the periods of data collection.

REFERENCES

- Chan, T.O., 2015. Cylindrical and polygonal object modelling and its use in lidar calibration and point cloud registration. Thesis, University of Calgary, AB, Canada.
- Coops, N.C., Goodbody, T.R.H., Cao, L., 2019. Four steps to extend drone use in research. *Nature*, 572(7770), 433-435.
- Cortes, C., Shahbazi, M., Ménard, P., 2018. UAV-LICAM system development: Calibration and geo-referencing. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-1, 107-114. doi.org/10.5194/isprs-archives-XLII-1-107-2018
- Ham, Y., Han, K.K., Lin, J.J., Golparvar-Fard, M., 2016. Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs): a review of related works. *Visualization in Engineering*, 4(1), 1.
- Kelly, J., Sukhatme, G.S., 2011. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor Self-calibration. *The International Journal of Robotics Research*, 30(1), 56-79.
- Kerle, N., Nex, F., Gerke, M., Duarte, D., Vetrivel, A., 2019. UAV-based structural damage mapping: A review. *ISPRS International Journal of Geo-Information*, 9(1), 14.
- Lupton, T., Sukkarieh, S., 2009. Efficient integration of inertial observations into visual SLAM without initialization. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1547-1552.
- Lynen, S., Achteleik, M.W., Weiss, S., Chli, M., Siegwart, R., 2013. A robust and modular multi-sensor fusion approach applied to MAV navigation. In *Proceedings of IEEE/RSJ international conference on intelligent robots and systems*, 3923-3929.
- Mirzaei, F.M., Roumeliotis, S.I., 2008. A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE transactions on robotics*, 24(5), 1143-1156.
- Nouwakpo, S.K., Wetz, M.A., McGwire, K., 2016. Assessing the performance of structure-from-motion photogrammetry and terrestrial LiDAR for reconstructing soil surface microtopography of naturally vegetated plots. *Earth Surface Processes and Landforms*, 41(3), 308-322.
- Ren, H., Zhao, Y., Xiao, W., Hu, Z., 2019. A review of UAV monitoring in mining areas: Current status and future perspectives. *International Journal of Coal Science & Technology*, 1-14.
- Shahbazi, M., Sohn, G., Théau, J., Ménard, P., 2017. Robust structure-from-motion computation: application to open-pit mine surveying from unmanned aerial images. *Journal of Unmanned Vehicle Systems*, 5(4), 126-145..
- Shahbazi, M., Sohn, G., Théau, J., Menard, P., Shahbazi, M., Sohn, G., Théau, J., Menard, P., 2015. Development and evaluation of a UAV-photogrammetry system for precise 3D environmental modeling. *Sensors*, 15(11), 27493-27524.
- Shahbazi, M., Théau, J., Ménard, P., 2014. Recent applications of unmanned aerial imagery in natural resource management. *GIScience & Remote Sensing*, 51(4), 339-365.
- Vázquez-Arellano, M., Griepentrog, H.W., Reiser, D., Paraforos, D.S., 2016. 3-D imaging systems for agricultural applications- A review. *Sensors*, 16(5), 618.
- Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR system with application to forest inventory. *Remote Sensing*, 4(6), 1519-1543
- Weiss, S., Siegwart, R., 2011. Real-time metric state estimation for modular vision-inertial systems. In *Proceedings of IEEE international conference on robotics and automation*, 4531-4537.
- Yao, H., Qin, R., Chen, X., 2019. Unmanned aerial vehicle for remote sensing applications- A review. *Remote Sensing*, 11(12), 1443.