COMPARATIVE EVALUATION OF DERIVED IMAGE AND LIDAR POINT CLOUDS FROM UAV-BASED MOBILE MAPPING SYSTEMS

Tian Zhou [1], Seyyed Meghdad Hasheminasab [1], Yi-Chun Lin [1], Ayman Habib [1], *

[1] Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907, USA - (zhou732, hashemin, lin934, ahabib)@purdue.edu

Commission II, WG II/2, ICWG I/II

KEY WORDS: UAVs, Image-based point cloud, structure from motion, dense matching, LiDAR

ABSTRACT:

Unmanned aerial vehicles (UAVs) have been widely used for 3D reconstruction/modelling in various applications such as precision agriculture, coastal monitoring, and emergency management. For such mapping applications, camera and LiDAR are the two most commonly used sensors. Mapping with imagery-based approaches is considered to be an economical and effective option and is often conducted using Structure from Motion (SfM) techniques where point clouds and orthophotos are generated. In addition to UAV photogrammetry, point clouds of the area of interest can also be directly derived from LiDAR sensors onboard UAVs equipped with global navigation satellite systems/inertial navigation systems (GNSS/INS). In this study, a custom-built UAV-based mobile mapping system is used to simultaneously collect imagery and LiDAR data. Derived LiDAR and image-based point clouds are investigated and compared in terms of their absolute and relative accuracy. Furthermore, stability of the system calibration parameters for the camera and LiDAR sensors are studied using temporal datasets. The results show that while LiDAR point clouds demonstrate a high absolute accuracy over time, image-based point clouds are not as accurate as LiDAR due to instability of the camera interior orientation parameters.

1. INTRODUCTION

Unmanned aerial vehicles (UAVs) equipped with global navigation satellite systems/inertial navigation systems (GNSS/INS) are becoming more popular for many applications because of their capability to carry advanced sensors and collect both high temporal and high spatial resolution data. UAV-based systems can provide accurate 3D spatial information at a relatively low cost, and therefore facilitate various applications including precision agriculture (Moghimi et al., 2020; Ravi et al., 2019; Masjedi et al., 2018; He et al., 2018; Habib et al., 2016), infrastructure monitoring (Greenwood et al., 2019), and archaeological documentation (Lin et al., 2019; Hamilton, Stephenson, 2016). RGB frame camera and LiDAR are the most common means to generate 3D point clouds for topographic mapping. Digital frame cameras onboard UAVs have been shown as a flexible and economical option for 3D reconstruction. The reconstructed image-based 3D model can be georeferenced using either ground control points (GCPs), known as indirect georeferencing, or trajectory information provided by a survey-grade GNSS/INS unit onboard the UAV, known as direct georeferencing. LiDAR-based systems, on the other hand, will directly lead to precise and high-resolution 3D point clouds. However, georeferencing of LiDAR point clouds coming from mobile mapping systems must be conducted using the direct georeferencing technique. System calibration of UAV-based GNSS/INS-assisted imaging and/or ranging systems is a vital step for direct georeferencing, and consequently reconstructing accurate LiDAR/image-based point clouds. System calibration parameters consist of internal characteristics of the onboard camera/LiDAR sensors, as well as mounting parameters relating the GNSS/INS body frame to the camera/LiDAR frames. When using direct georeferencing, any deviation in the system

calibration parameters from their true values will adversely affect the accuracy of reconstructed object space.

Point clouds coming from LiDAR and imagery have been evaluated and compared in different studies. Ni et al. (2014) studied the possibility of using image-based points clouds instead of LiDAR data in forested areas for biomass estimation. The authors analyzed and compared the image-based point cloud data to small footprint LiDAR data and large footprint LiDAR waveform data. They showed that the satellite stereo imagery could result in point clouds with points on both canopy and ground surfaces in unclosed forest but only the canopy surface in dense forest. Hence, in such cases, to calculate the height of the canopy using imagery, terrain elevation from other sources would be needed. Thiel and Schmullius (2017) compared UAV image-based point clouds and manned airborne LiDAR data over a forested area. Their results showed a high correlation between LiDAR and image-based point clouds with a slight superiority of the results coming from the latter. More specifically, while using LiDAR data, 45 out of 205 trees were not detected, and this number was only 14 when UAV image data was used. Elsner et al. (2018) compared UAV image-based 3D reconstruction results with wheel-based mobile LiDAR and manned airborne LiDAR. Their results suggested that image-based point cloud has consistently higher elevation than LiDAR data for all the utilized datasets. UAV-based photogrammetry and LiDAR were compared in a study conducted by Shaw et al. (2019). Similar to the results shown by Elsner et al. (2018), image-based point cloud showed a constant positive elevation bias from 4 to 9 cm when compared to the LiDAR surfaces and Real-Time Kinematic (RTK)-GNSS measurements. Lin et al. (2019) evaluated the relative performance of UAV LiDAR in mapping coastal environment when compared to the UAV photogrammetry. Their

---

* Corresponding author

results suggested that both LiDAR and image-based point clouds had a good degree of alignment with an overall precision of ±5 to ±10 cm, with the LiDAR data outperforming the photogrammetric surface in terms of point density, ground coverage and ability to penetrate through vegetation.

Despite the fact that the discrepancy between LiDAR and image-based point clouds has been widely observed and reported, to the best of the authors' knowledge, there is no investigation regarding the cause of such differences between the two sets of point clouds. To address this issue, this work aims at exploring possible factors that can cause discrepancies between LiDAR and image-based point clouds by using temporal datasets with different sensor settings. Temporal datasets are used to evaluate system calibration parameters, while the impact of sensor settings on the reconstructed point clouds is investigated through changes in camera focus settings.

In this paper, image-based sparse point cloud is generated through a GNSS/INS-assisted Structure from Motion (SfM) strategy introduced by Hasheminasab et al. (2020). Then, dense point cloud is generated using an approach similar to the patch-based multi-view stereo (PMVS) algorithm (Furukawa, Ponce, 2009). LiDAR point cloud is reconstructed with the help of GNSS/INS trajectory and system calibration parameters. To evaluate the absolute accuracy of the derived point clouds, checkerboard targets are deployed in the study site and surveyed using RTK-GNSS technique. Also, a point-to-point strategy is proposed for assessing relative accuracy between the two sets of point clouds.

The remainder of the paper is organized as follows: Section 2 introduces the UAV-based data acquisition system and the datasets used in this study, Section 3 describes the approaches for deriving image-based and LiDAR-based point clouds as well as the strategies for assessing the quality of the point clouds, Section 4 presents the experimental results, and Section 5 provides conclusions and recommendations for future work.

2. DATA ACQUISITION SYSTEM SPECIFICATIONS AND DATASETS DESCRIPTION

In this study, LiDAR and image-based point clouds are generated from data captured by a custom-built UAV-based mobile mapping system, shown in Figure 1. The system consists of a Dà-Jiāng Innovations (DJI) Matrice 600 Pro (M600P) carrying a Sony α7R III (ILCE-7RM3) RGB camera, a Velodyne Puck Lite LiDAR sensor, and Applanix APX-15 UAV v3 GNSS/INS unit. The LiDAR unit, RGB camera, and GNSS/INS unit are rigidly fixed to one another. Direct geo-refencing information, i.e., the position and orientation of the system, is provided by the APX-15 v3 unit at 200 Hz data rate. After post-processing of the GNSS/INS data, the expected positional accuracy is 2-5 cm, and the accuracy for pitch/roll and heading is 0.025° and 0.08°, respectively (Applanix, 2020). The Sony α7R III camera is a 42-megapixel camera with a 7952 × 5304 complementary metal oxide semiconductor (CMOS) array, 4.5 μm pixel size, and a lens with 35 mm nominal focal length (Sony, 2020). The Velodyne Puck LITE, with a weight of 590 g, is a lighter version of VLP-16 Puck (830 g) and consists of 16 channels. This sensor generates approximately 300,000 points per second with a 360° horizontal field of view and a 30° vertical field of view (±15° from the horizon). The maximum measurement range is 100 m with a ±3 cm range accuracy (Velodyne, 2020).
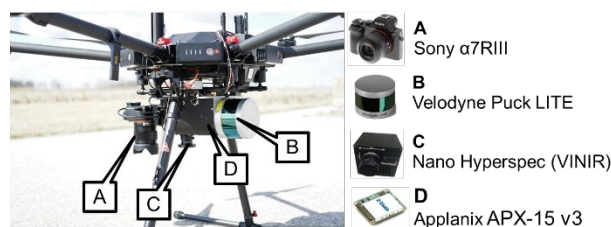

Figure 1. The UAV-based mobile mapping system and onboard sensors used in this study.

As mentioned earlier, rigorous system calibration is essential for achieving accurate georeferenced products through direct georeferencing. In this study, camera IOPs are estimated using the United States Geological Survey (USGS) simultaneous multi-frame analytical calibration (SMAC) distortion model through an indoor calibration procedure similar to the one proposed by He and Habib (2015). Derived square root of a posteriori variance of the indoor calibration bundle adjustment procedure is 0.73 pixel, which is an indication of high precision of estimated IOPs. Mounting parameters, i.e., boresight angles and lever arm components, between GNSS/INS unit and the onboard imaging/ranging sensors are estimated through the rigorous system calibration procedure proposed by Ravi et al. (2018) using a calibration dataset collected on November 11th, 2019. Figure 2 depicts the qualitative evaluation of the system calibration by showing the alignment among LiDAR, image-derived, and ground control points of a checkerboard target, where the planimetric and vertical alignments are shown by top view and slide view, respectively. As shown in Figure 2, the absolute accuracy of the LiDAR and image-based points is in the range of ±2 to ±5 cm compared to the established GCP.
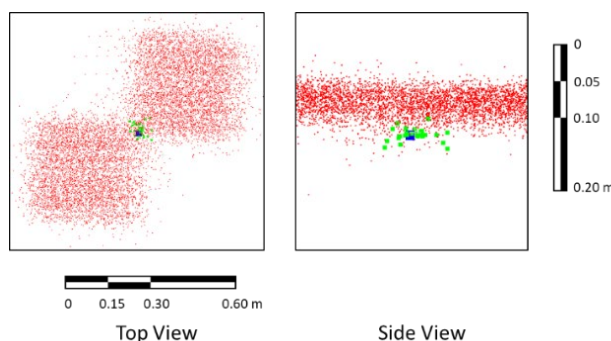

Figure 2. Qualitative evaluation of 3D alignment among LiDAR (red), image points (green), and GCP (blue) after calibration for checkerboard target.

In this study, four datasets were collected over a study site at Purdue's Agronomy Center for Research and Education (ACRE) with different geomorphic features, i.e., grass, pavement, and building roof. To study the impact of camera settings on the reconstructed image-based point cloud, different focus settings were selected for the Sony camera, i.e., auto focus and manual focus. Table 1 summarizes the flight configurations and camera settings used for different datasets. The ground sampling distance (GSD) of the imagery is 0.6 cm at 41 m flying height. As reported in Table 1, auto focus and manual focus modes were used in datasets A and B, respectively.

To evaluate the absolute accuracy of the reconstructed point clouds, a total of twelve highly reflective checkboard targets were deployed in the study site, as shown in Figure 3. Centers of these

checkerboard targets were surveyed through the RTK-GNSS technique using a Trimble R10 GNSS receiver. Considering the distance between the rover receiver and the GNSS base-station, i.e., 6 km, the expected horizontal and vertical accuracy from the R10 is in the range of ±2 to ±3 cm and ±3 to ±4 cm, respectively.

Dataset Name	Date	Camera Focus Settings	Flying Height (m)	Ground Speed (m/s)	Lateral Distance (m)	Imagery Overlap/ Side-Lap (%)
A-1	20200304	Auto focus	41	4.0	6	83/89
A-2	20200501	Auto focus	41	4.0	6	83/89
B-1	20200304	Manual focus	41	4.0	6	83/89
B-2	20200501	Manual focus	41	4.0	6	83/89

Table 1. Flight configurations and camera focus settings for the flight missions captured for the experimental datasets
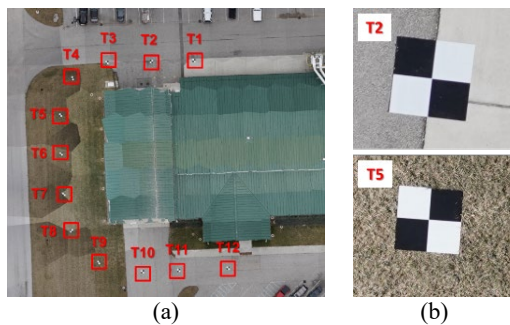

Figure 3. (a) Study site and deployed checkboard targets highlighted with the red boxes, and (b) sample images of the targets.

3. METHODOLOGY

In this section, first, the approaches for image-based and LiDAR-based point cloud generation are described. Then, the strategies for accuracy evaluation of the derived 3D point clouds are discussed.

3.1 Image-based Point Cloud Generation

In this study, image-based point clouds are generated through five steps, namely, stereo-image matching, relative orientation parameter (ROP) estimation, exterior orientation parameter (EOP) recovery, bundle adjustment (BA), and dense matching.

In stereo-image matching step, Scale Invariant Feature Transform (SIFT) (Lowe, 2004) algorithm is used to detect local features along with their descriptors. Then, similar to what was introduced by Hasheminasab et al. (2020), rather than conducting a traditional exhaustive search among the feature descriptors, the available GNSS/INS trajectory is used to reduce the search space and consequently the matching ambiguity. Matching ambiguity is common when dealing with images captured over areas with homogeneous nature, which is the case for the dataset used in this study consisting of homogeneous patches of grass, pavement, and building roofs. Once conjugate features are established between overlapping images, the relative orientation parameters including two positional and three rotational parameters are estimated between the stereo-pairs using coplanarity constraint.

This constraint enforces the coplanarity of the light rays connecting the perspective centers of the imaging sensor, object point, and the respective image points. In the EOP recovery step, a local coordinate system is first defined by a stereo-pair that has the maximum number of feature correspondences. The remaining images are then sequentially augmented into this reference frame through rotation and translation averaging techniques. Then, a GNSS/INS-assisted bundle adjustment is conducted to refine the derived EOPs, 3D coordinates of the object points, and boresight angles. These four steps together comprise the Structure-from-Motion (SfM) framework. Figure 4 shows a sample of generated sparse point cloud and corresponding orthophoto from A-1 dataset. As can be seen in Figure 4, the reconstructed image-based 3D points (30,000 points) are not well-distributed over the study site, i.e., most points belong to grass areas while few points are reconstructed on building roof and pavement surfaces.
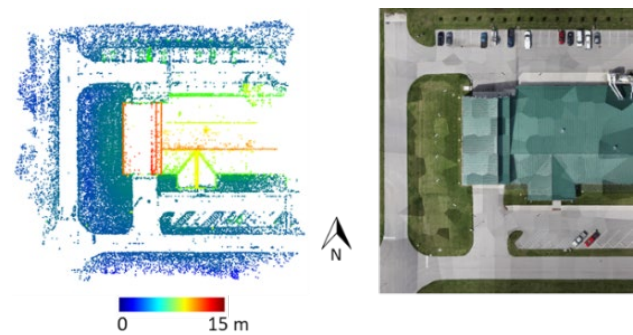

Figure 4. Generated sparse point cloud (colored by height) with 30,000 points and corresponding orthophoto for A-1 dataset.

In order to generate well-distributed dense point clouds, a dense matching strategy inspired by the PMVS algorithm (Furukawa, Ponce, 2009) is implemented. The approach starts with detecting Harris features (Harris, Stephens, 1988) in each image. To conduct feature matching, each image in the image block is sequentially selected as reference, and its K-nearest images (K=10 in this study) are considered as candidate images. Using the refined EOPs from SfM results, the epipolar constraint is applied to establish candidate matching features between the reference image and each candidate image. Next, a patch corresponding to a given pair of matching features is defined as follows: a) the 3D point derived by light-ray intersection from the matching features is regarded as center; b) the norm of the patch is defined so that it is oriented toward the reference image; and c) size of patch is pre-defined according to GSD of the image, e.g., 6 cm × 6 cm patches. Then, corresponding patches in image-space are derived through back-projection of the patch in question into all candidate images. In the next step, Normalized Cross Correlation (NCC) is used to measure the similarity between image patches in reference image and each candidate image. A patch in the reference image is retained if at least five candidate images are matched. Going through all the images, a sparse set of patches is generated in the object space. And then a denser set of patches is obtained through an expansion/interpolation procedure. The final dense point cloud is derived after visibility and consistency checks are conducted within the point clouds where wrong and/or occluded patches are removed. Figure 5 depicts a sample of generated dense point cloud from A-1 dataset. Comparing Figure 5 with Figure 4, one can easily observe the improvement in the distribution as well as the number of points in the generated dense point cloud.
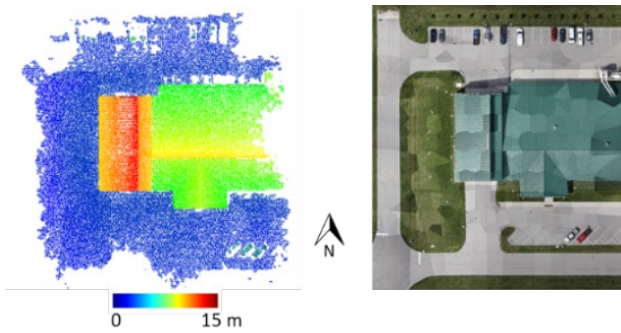
Figure 5. Generated dense point cloud (colored by height) with 80,000 points and corresponding orthophoto for A-1 dataset.

3.2 LiDAR Point Cloud Generation

Raw LiDAR data includes intensity and range measurements along with the direction where the laser beams is pointing. By coupling the range and orientation of the laser beams, we can obtain the position of the laser beam footprint relative to the laser unit frame, denoted by $r_I^{lu}(t)$. The position, $r_v^m(t)$, and orientation, $R_v^m(t)$, of the vehicle frame relative to the mapping frame are derived through direct georeferencing. The laser unit frame and vehicle frame are related using the calibrated mounting parameters, which consist of the lever arm, r_{lu}^v, and the boresight rotation matrix, R_{lu}^v. The 3D coordinates of a ground point, I, in the mapping frame can then be derived using Equation 1 (El-Sheimy et al., 2005).

$$r_I^m = r_v^m(t) + R_v^m(t)r_{lu}^v + R_v^m(t)R_{lu}^v r_I^{lu}(t) \qquad (1)$$

In this study, the rotation axis of the Velodyne onboard the UAV is parallel to the flying direction. To ensure that the LiDAR points are mainly coming from the object space in question, i.e., area below UAV, an object point is reconstructed only when the direction of corresponding laser beam is less than ±70° from nadir. Reconstructed LiDAR point cloud from A-1 dataset is illustrated in Figure 6.
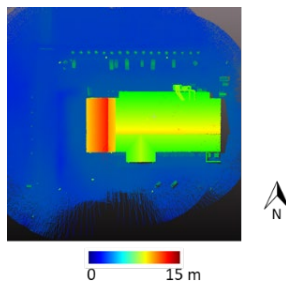

Figure 6. Reconstructed LiDAR point cloud (colored by height) with 31,000,000 points for A-1 dataset.

3.3 Accuracy and Comparative Analysis of Generated Point Clouds

In this section, we first present the strategies for evaluating the absolute accuracy of the derived LiDAR and image-based point clouds against RTK-GNSS survey. The approach for estimating the relative accuracy between the LiDAR and image-based point clouds is then introduced.

The absolute accuracy of the derived point cloud is assessed against the RTK-GNSS measurements of twelve checkerboard

targets that were set up in the field before data acquisition. Starting with establishing the point correspondence between the RTK-GNSS points and the LiDAR and image-based point cloud, the coordinate differences between the point pairs are calculated, and the statistics including the mean, standard deviation (STD), and root-mean-square-error (RMSE) are reported. Conducting a reliable comparison between points (RTK-GNSS measurements) and point clouds is contingent on establishing point correspondence, which is introduced below:

- *Image-based point cloud*: The center points of the checkerboard targets are manually identified in the images. The 3D coordinates of the center points, hereafter denoted as image-derived 3D coordinates of targets, are then estimated through multi light-ray intersection, using camera IOPs and the refined EOPs derived from the SfM strategy.

- *LiDAR-based point cloud*: In the first step, centers of the highly reflective checkerboard targets are manually identified from the LiDAR point cloud based on the intensity information, and denoted as initial points. The initial points are expected to have a horizontal accuracy of ±3 to ±5 cm due to the noise level of the LiDAR data caused by i) the GNSS/INS trajectory errors, and ii) the nature of LiDAR pulse returns from highly reflective surfaces. Then, a strategy based on iterative plane fitting is proposed in order to derive reliable Z coordinates. First, a spherical region centered at each initial point is created with a pre-defined radius, e.g., 0.5 m. An iterative plane fitting is conducted using all the LiDAR points in the spherical region to find the best fitted plane. Finally, the center point is defined as the projection of the initial point on the best-fitted plane, and hereafter denoted as LiDAR-derived 3D coordinates of the targets. Figure 7 illustrates a sample of deriving point correspondence in LiDAR point cloud. Figure 7(a) shows the RGB image of a highly reflective checkerboard target, where the coordinates of the target center are determined through an RTK-GNSS survey. Figure 7(b) depicts a sample of highly reflective checkerboard target in the LiDAR point cloud, the initial point (red point), and the LiDAR-derived 3D coordinates of the targets (green point).
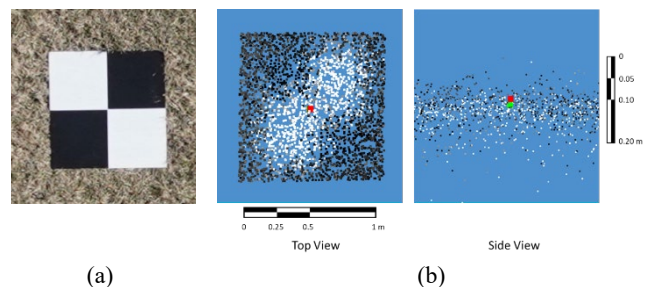

(a)   (b)

Figure 7. A sample of point correspondence in LiDAR point cloud: (a) a checkerboard target in an RGB image, and (b) the LiDAR points of the checkerboard target, initial point, (red point), and LiDAR-derived center of the target (green point). The LiDAR point cloud is colored by intensity, where points with low and high intensities are shown in black and white, respectively.

In addition to evaluating absolute accuracy of the point clouds, relative accuracy between the LiDAR and image-based point clouds is assessed. The differences between the LiDAR and image-based point clouds are evaluated on a point-by-point basis. The point correspondence between two sets of point clouds is first established. For a given point in image-based point cloud (denoted as image point), its closest LiDAR point is identified. To ensure reliable correspondence, the distance between these two points must be smaller than a pre-defined radius, e.g., 1 m. Next, a spherical region centered at the source point with a pre-

defined threshold, e.g., 0.5 m, is created. Then, an iterative plane fitting is conducted using all the LiDAR points in the spherical region. The plane is considered valid if the RMSE of normal distances of the points from the best-fitting plane is smaller than a pre-defined threshold, e.g., 0.3 m, and ratio of retained LiDAR points is more than 50%. Finally, the image point is projected onto the best-fitted plane, and the projection defines the corresponding LiDAR point. Once the point correspondence is established, the coordinate differences between the image points and the corresponding LiDAR points can be evaluated.

One should note that due to the homogeneous texture of the gable roof in the study site, the majority of reconstructed points are on the ground surface with horizontal orientation. Therefore, discrepancy between LiDAR and image-based point clouds can be reliably evaluated only in vertical direction. Therefore, only Z discrepancies between the LiDAR and image-based point clouds will be considered as the criterion for relative accuracy evaluation.

4. EXPERIMENTAL RESULTS

This section reports the results of absolute and relative accuracy of the LiDAR and image-based point clouds for the four datasets. As mentioned in Section 2, LiDAR and image-based point clouds are generated using system calibration parameters estimated from a dataset collected on November 11th, 2019.

4.1 Absolute Accuracy of Image-based Point Cloud

As described in Section 3.3, image-derived 3D coordinates of the targets are estimated through multi light-ray intersection. The mean, STD, and RMSE of the differences between the image-based and RTK-GNSS coordinates for the twelve targets are reported in Table 2. According to the statistics reported in Table 2, there is a misalignment between image-based point clouds and RTK-GNSS measurements of the targets in all X, Y, and Z directions. Moreover, one can observe that large Z RMSE errors are mainly caused by a constant shift between image and target points along vertical direction (large mean Z values in Table 2). On the other hand, large STD values in X and Y directions show a horizontal misalignment of the image-based point cloud. Figures 8(a) and 8(b) depict the X and Y differences between the image-based and RTK-GNSS coordinates for the twelve targets for A-1 dataset, respectively. As shown in Figure 8, there is a systematic difference in both X and Y coordinates. More specifically, the differences increase as the points deviate further away from the mean (represented by blue line in Figure 8). Based on the large STD reported in Table 2 and the pattern of differences shown in Figure 8, one can conclude that there is a scaling issue in the XY-plane, which is caused by inaccurate system calibration parameters. Considering the fact that inaccurate lever-arm components only cause constant shifts in the object space and boresight angles are solved for during the bundle adjustment process, it can be deduced that non-optimal camera IOPs are the source of error. Moreover, since the principal distance only leads to variation along the Z direction, the inaccurate distortion parameters are most likely to be the cause of the scaling issue. In conclusion, although accurate IOPs were estimated in the system calibration procedure, interior camera parameters are hypothesized to be changing over time. This is also observed by comparing the derived absolute accuracy of datasets which have the same camera settings but captured on different dates, i.e., A-1/A-2 with auto focus settings and B-1/B-2 with manual focus settings. Looking again into Table 2, one can note that RMSE of differences between image-based and RTK-GNSS coordinates of the targets changes from 6-10 cm in

A-1 dataset collected on March 2020, to 12-27 cm in A-2 dataset collected on May 2020. Similar results hold true for datasets with manual focus settings, i.e., B-1/B-2. In addition, comparing the results of A-1/B-1 and A-2/B-2 datasets, the accuracy of image-based point clouds is not consistent using different focus modes, which further proves that the utilized focus mode can affect the accuracy of the derived point cloud.

Dataset	Statistics Criteria	X Differences (m)	Y Differences (m)	Z Differences (m)
A-1	Mean(m)	0.04	0.02	0.10
	STD (m)	0.04	0.07	0.03
	RMSE (m)	0.06	0.07	0.10
A-2	Mean(m)	0.08	0.01	0.27
	STD (m)	0.10	0.16	0.02
	RMSE (m)	0.12	0.16	0.27
B-1	Mean(m)	0.03	0.00	0.13
	STD (m)	0.03	0.05	0.02
	RMSE (m)	0.04	0.05	0.13
B-2	Mean(m)	0.05	0.01	0.23
	STD (m)	0.07	0.11	0.03
	RMSE (m)	0.08	0.11	0.23

Table 2. Mean, STD, and RMSE of the differences between image-derived and RTK-GNSS coordinates of the twelve check points for the four datasets.
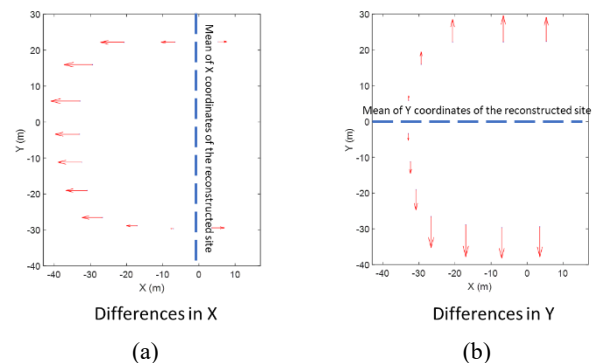

Figure 8. Differences in (a) X coordinates and (b) Y coordinates shown by red arrows between image-derived and RTK-GNSS coordinates of the twelve ground targets for A-1 dataset where blue lines represent mean of X and Y coordinates of the reconstructed site (magnitude of the arrows is magnified by a factor of 100).

4.2 Absolute Accuracy of LiDAR Point Cloud

To assess the absolute accuracy of the LiDAR point cloud, LiDAR-derived 3D coordinates of the twelve checkerboards were computed using the strategy proposed in Section 3.3. The absolute accuracy is assessed by evaluating the differences between the LiDAR-derived coordinates and the corresponding RTK-GNSS.

Figure 9 shows the LiDAR-based point cloud colored by intensity and the RTK-GNSS measurements (the red points) of two ground targets in A-1 dataset. As can be seen from the top and side view illustrations in Figure 9, the LiDAR and RTK-GNSS points are well-aligned in the horizontal and vertical directions. Quantitative accuracy assessment results for the four datasets are listed in Table 3. According to Table 3, the RMSE values of differences in X and Y directions for all datasets are

within 5 cm. One should note that these planimetric differences arise from the difficulty in identifying the actual center of targets in the LiDAR data. Also, looking into Z differences in Table 3, RMSE values in the vertical direction are in the range of 2-4 cm. Overall, small RMSE values reported in Table 3 verify the accuracy and reliability of the LiDAR data over two different dates. Furthermore, the compatibility between LiDAR and RTK-GNSS validates the stability of system parameters regarding the LiDAR sensor.
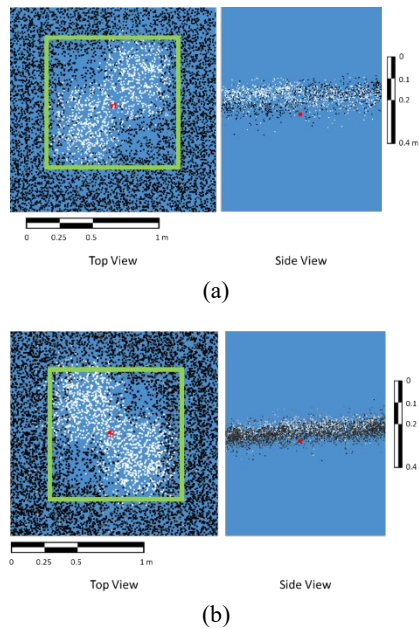

(a)

(b)

Figure 9. Alignment between LiDAR data (colored by intensity) and RTK-GNSS points (colored in red) of checkerboard targets: (a) Target 3 and (b) Target 9 for A-1 dataset.

Dataset	Statistics Criteria	X Differences (m)	Y Differences (m)	Z Differences (m)
A-1	Mean(m)	0.03	0.01	0.03
	STD (m)	0.02	0.03	0.01
	RMSE (m)	0.04	0.03	0.04
A-2	Mean(m)	-0.01	0.00	0.01
	STD (m)	0.03	0.02	0.02
	RMSE (m)	0.03	0.02	0.02
B-1	Mean(m)	-0.02	0.01	0.03
	STD (m)	0.03	0.05	0.02
	RMSE (m)	0.03	0.05	0.04
B-2	Mean(m)	-0.01	0.00	0.01
	STD (m)	0.03	0.03	0.02
	RMSE (m)	0.03	0.03	0.02

Table 3. Mean, STD, and RMSE of the differences between LiDAR-derived and RTK-GNSS coordinates of the twelve check points for the four datasets.

4.3 Comparative Quality Assessment of Image-based and LiDAR Point Clouds

In this section, a comparison between LiDAR and image-based point clouds is conducted for the four datasets. As described in Section 3, for each point in image-derived dense point cloud, corresponding LiDAR point is identified using the point-to-plane matching technique. To evaluate the relative accuracy between the two sets of point clouds, only Z discrepancy is considered due

to the fact that the majority of reconstructed points belong to horizontal surfaces in the study site. Figure 10(a) shows the image-based point cloud for A-2 dataset, where each point is colored by its height difference from the corresponding LiDAR point, ranging from +15 cm (blue) to +30 cm (red). Also, in order to have a better visualization of the derived image-based point cloud for A-2 dataset, each point is colored by its height value and is illustrated in Figure 10(b), ranging from 0 m (blue) to 15 m (red). In addition, an orthophoto of the A-2 dataset is shown in Figure 10(c). Looking into Figure 10(a) one can observe spatial discrepancy patterns over the study site. More specifically, the elevation difference for ground points changes from +30 cm to +15 cm from central part to the perimeter of the study site. Moreover, the elevation differences between the point clouds depend on the height of the reconstructed object points. For points on the highest part of the roof (western part, see Figure 10(b)), elevation differences are approximately +15 cm, while points in neighboring ground area exhibits +30 cm vertical discrepancy from LiDAR points. For other three datasets, i.e., A-1, B-1, and B-2, similar discrepancy patterns are found, which indicate that the Z discrepancy between the two sets of point clouds is not coming from random errors, but caused by non-optimal system calibration parameters. Similar to what was explained in Section 4.1, it can be concluded that the camera IOPs are the source of the observed scaling issue in the image-based point clouds.
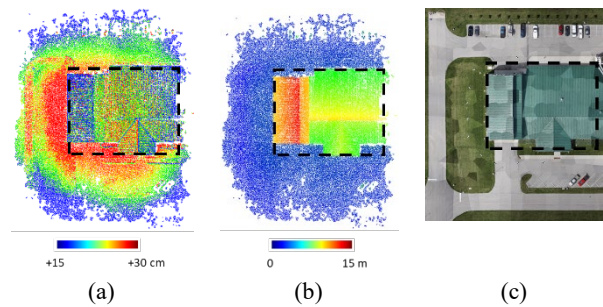

(a)   (b)   (c)

Figure 10. Image-based point cloud colored by (a) Z discrepancy to LiDAR point cloud and (b) height, as well as (c) orthophoto for A-2 dataset.

5. CONCLUSIONS AND RECOMMENDATIONS FOR RFUTURE WORK

In this study, a comprehensive comparison between UAV-derived LiDAR and image-based point clouds has been conducted in order to assess the potential of using UAVs for 3D mapping applications. The experimental results show that the LiDAR point clouds are accurate and reliable over time, which further proves the stability of LiDAR system calibration parameters. However, it has been observed that there is a variation in camera IOPs over time, which adversely affects the accuracy of the derived point cloud. Considering the fact that consumer grade cameras like the Sony α7R III used in this study are not designed specifically for mapping applications, variation in their internal characteristics over time should be expected. Hence, frequent camera calibration needs to be done for image-based mapping applications that require high accuracy. In this regard, using LiDAR-derived control points for refining camera IOPs will be a possible focus of the future work. Consequently, camera IOPs can be refined in every flight for UAV-based mobile mapping systems equipped with LiDAR sensors.

ACKNOWLEDGEMENTS

REFERENCES

Applanix. Apx-15 UAV Datasheet. Available online: https://www.applanix.com/downloads/products/specs/APX15_UAV.pdf (accessed on 8 March 2020).

El-Sheimy, N., Valeo, C., & Habib, A. (2005). Digital terrain modeling: acquisition, manipulation, and applications. *Artech House*.

Elsner, P., Dornbusch, U., Thomas, I., Amos, D., Bovington, J., & Horn, D. (2018). Coincident beach surveys using UAS, vehicle mounted and airborne laser scanner: Point cloud inter-comparison and effects of surface type heterogeneity on elevation accuracies. Remote Sensing of Environment, 208, 15-26.

Furukawa, Y., & Ponce, J. Accurate, dense, and robust multiview stereopsis. *IEEE T. Pattern Anal*. 2009, 32(8), 1362-1376.
Greenwood, W. W., Lynch, J. P., & Zekkos, D. (2019). Applications of UAVs in civil infrastructure. *Journal of Infrastructure Systems*, *25*(2), 04019002.

Habib, A., Xiong, W., He, F., Yang, H. L., & Crawford, M. (2016). Improving orthorectification of UAV-based push-broom scanner imagery using derived orthophotos from frame cameras. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(1), 262-276.

Harris, C. G., & Stephens, M. (1988, August). A combined corner and edge detector. In *Alvey vision conference* (Vol. 15, No. 50, pp. 10-5244).

Hasheminasab, S. M., Zhou, T., & Habib, A. (2020). GNSS/INS-Assisted Structure from Motion Strategies for UAV-Based Imagery over Mechanized Agricultural Fields. *Remote Sensing*, 12(3), 351.

Hamilton, S., & Stephenson, J. (2016). Testing UAV (drone) aerial photography and photogrammetry for archeology. Lakehead University, *Tech. Rep*.

He, F., & Habib, A. (2015,). Target-based and Feature-based Calibration of Low-cost Digital Cameras with Large Field-of-view. In *ASPRS 2015 Annual Conference, Tampa, Florida*.

He, F., Zhou, T., Xiong, W., Hasheminnasab, S. M., & Habib, A. (2018). Automated aerial triangulation for UAV-based mapping. *Remote Sensing*, *10*(12), 1952.

Lin, Y. C., Cheng, Y. T., Zhou, T., Ravi, R., Hasheminasab, S. M., Flatt, J. E., ... & Habib, A. (2019). Evaluation of UAV LiDAR for Mapping Coastal Environments. *Remote Sensing*, 11(24), 2893.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), 91-110.

Masjedi, A., Zhao, J., Thompson, A. M., Yang, K. W., Flatt, J. E., Crawford, M. M., ... & Chapman, S. (2018, July). Sorghum biomass prediction using UAV-based remote sensing data and crop model simulation. *In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 7719-7722). IEEE.

Moghimi, A., Yang, C., & Anderson, J. A. (2020). Aerial hyperspectral imagery and deep neural networks for high-throughput yield phenotyping in wheat. *Computers and Electronics in Agriculture*, 172, 105299.

Ni, W., Ranson, K. J., Zhang, Z., & Sun, G. (2014). Features of point clouds synthesized from multi-view ALOS/PRISM data and comparisons with LiDAR data in forested areas. *Remote sensing of environment*, 149, 47-57.

Ravi, R., Lin, Y. J., Elbahnasawy, M., Shamseldin, T., & Habib, A. (2018). Simultaneous system calibration of a multi-lidar multicamera mobile mapping platform. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 11(5), 1694-1714.

Ravi, R., Hasheminasab, S. M., Zhou, T., Masjedi, A., Quijano, K., Flatt, J. E., ... & Habib, A. (2019, May). UAV-based multi-sensor multi-platform integration for high throughput phenotyping. In *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IV* (Vol. 11008, p. 110080E). International Society for Optics and Photonics.

Shaw, L., Helmholz, P., Belton, D., & Addy, N. (2019). COMPARISON OF UAV LIDAR AND IMAGERY FOR BEACH MONITORING. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Sony. α7R III Full Specifications and Features. Available online: https://www.sony.com/electronics/interchangeable-lens-cameras/ilce-7m3-body-kit/specifications (accessed on 8 March 2020).

Thiel, C., & Schmullius, C. (2017). Comparison of UAV photograph-based and airborne lidar-based point clouds over forest from a forestry application perspective. International journal of remote sensing, 38(8-10), 2411-2426.

Velodyne. Puck LITE Data Sheet. Available online: http://www.mapix.com/wp-content/uploads/2018/07/63-9286_Rev-H_Puck-LITE_Datasheet_Web.pdf (accessed on 8 March 2020).