Linear Feature-Based Image/LiDAR Integration for a Stockpile Monitoring and Reporting Technology

Seyyed Meghdad Hasheminasab, Tian Zhou, and Ayman Habib

***Abstract*—Stockpile monitoring has been recently conducted with the help of modern remote sensing techniques—e.g., terrestrial/aerial photogrammetry/LiDAR—that can efficiently produce accurate 3-D models for the area of interest. However, monitoring of indoor stockpiles still requires more investigation due to unfavorable conditions in these environments such as a lack of global navigation satellite system signals and/or homogenous texture. This article develops a fully automated image/LiDAR integration framework that is capable of generating accurate 3-D models with color information for stockpiles under challenging environmental conditions. The derived colorized 3-D point cloud can be subsequently used for volume estimation and visual inspection of stockpiles. The main contribution of the developed strategy is using automatically derived conjugate image/LiDAR linear features for simultaneous registration and camera/LiDAR system calibration. Data for this article are acquired using a camera-assisted LiDAR mapping platform—denoted as stockpile monitoring and reporting technology—which was recently designed as a time-efficient and cost-effective bulk material tracking. Experimental results on three datasets show that the developed framework outperforms a classical planar feature-based registration technique in terms of the alignment of acquired point cloud. Results also indicate that the proposed approach can lead to a high relative accuracy between image lines and their corresponding back-projected LiDAR features in the range of 4–7 pixels.**

***Index Terms*—Image/LiDAR integration, linear features, matching, point cloud colorization, point cloud registration, stockpile monitoring, system calibration.**

I. Introduction

STOCKPILE monitoring is important for managing bulk materials in the agriculture, construction, and mining industry. For transportation roadway maintenance in particular, accurate monitoring of salt stockpiles is important to improve safety and traffic flow during snowstorms. Thanks to recent advances in sensor and platform technologies, photogrammetric and LiDAR approaches are commonly adopted for stockpile monitoring. Photogrammetric approaches use an excessive number of overlapping images to produce 3-D point clouds—with color information—through structure from motion (SfM) [1], [2] and multiview stereo [3], [4] techniques. The main limitation of these approaches is their inability to find sufficient number of conjugate points when dealing with images exhibiting repetitive patterns and/or significant viewpoint change.

Different from image-based mapping techniques, LiDAR-based approaches can directly provide dense 3-D point cloud but without any spectral/color information. LiDAR sensors are either mounted on mobile mapping systems such as unmanned aerial vehicles or static tripods, e.g., terrestrial laser scanners (TLS). Mobile LiDAR systems depend heavily on an integrated global navigation satellite system (GNSS)/inertial navigation system for position and orientation estimation. Thus, their application is limited to open-sky areas. On the other hand, TLS systems can acquire high-resolution point clouds with millimeter-level precision without the need for GNSS signals. However, they have not been widely used for stockpile monitoring due to their high-cost and sometimes time-consuming data acquisition/postprocessing steps when dealing with large point clouds/facilities.

In order to overcome the limitations of current stockpile monitoring technologies, a camera-assisted LiDAR mapping platform denoted as *stockpile monitoring and reporting technology* (SMART) was designed and introduced in our previous work [5]. Different from system-driven technologies such as TLS that use sophisticated and expensive encoders, the SMART system focuses on a data-driven strategy by acquiring data in a simple, cost-effective procedure. The ultimate goal of the SMART system is to produce a well-aligned colorized point cloud that can be used for a realistic 3-D visualization and accurate volume estimation of stockpiles.

To derive well-aligned colorized point clouds, an image/LiDAR integration process is required. More specifically, aligning data from different sensors requires establishing the internal characteristics of individual sensors—known as interior orientation parameters (IOP), mounting parameters (lever-arm components and boresight angles) relating the different sensors, and registration parameters (if multiple point clouds defined relative to different coordinate frames are available). While LiDAR IOP—often provided by the manufacturer—are accurate and stable, consumer-grade cameras require frequent calibration due to the instability of their IOP [6]. Mounting parameters define the positional and rotational offsets between different sensors and are derived through a system calibration procedure. The frequency of the required system calibration depends on how rigidly the sensors are fixed with respect to each other. In addition, the registration process aims at estimating the 3-D rigid transformation parameters—i.e., three rotation angles and three

Manuscript received 5 September 2022; revised 22 January 2023; accepted 18 February 2023. Date of publication 28 February 2023; date of current version 16 March 2023. This work was supported in part by the Joint Transportation Research Program administered by the Indiana Department of Transportation under Grant SPR-4549 and in part by Purdue University. *(Corresponding author: Seyyed Meghdad Hasheminasab.)*

The authors are with the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: hashemin@purdue.edu; zhou732@purdue.edu; ahabib@purdue.edu).

Digital Object Identifier 10.1109/JSTARS.2023.3250392

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

translations (scale is considered as unity for a well-calibrated LiDAR)—between two point clouds. Finally, system calibration and registration parameters are used to derive the registered colorized point cloud.

Over the past few years, several studies have addressed the problem of image/LiDAR integration by focusing on system calibration. Camera/LiDAR calibration techniques estimate the system parameters by minimizing the discrepancy between conjugate features that are extracted from both sensors at the same or different locations. Depending on the type of utilized features, calibration techniques can be categorized into target-based and target-less approaches. In an early target-based calibration study, Zhang and Pless [7] used a planar checkerboard for establishing the mounting parameters between a camera and a 2-D LiDAR. They parameterized the checkerboard plane by its unit normal vector and its distance to the camera frame. Then, the camera/LiDAR mounting parameters were derived by minimizing LiDAR point to checkerboard plane distance in a nonlinear optimization procedure. Several studies extended this work for calibrating systems equipped with a camera and 3-D LiDAR sensors [8], [9], [10], [11]. In addition to point-to-plane geometric constraints, other feature correspondences such as line-to-plane [12] and point-to-point [13], [14] have been also adopted through designing custom-built targets, e.g., planar targets containing symmetric holes. Although target-based strategies produce accurate results, using these approaches might not be always practical for applications where calibration parameters change from one mission to another.

Early works dealing with *in-situ* calibration, using target-less strategies, were based on manual identification of conjugate natural point and linear features in indoor scenes [15], [16]. Several efforts have been made toward developing fully automated camera/LiDAR calibration frameworks. The majority of these techniques—also referred to as motion-based approaches—use visual-odometry [17], [18] or SfM [19], [20] to establish conjugate features between image and LiDAR data. More specifically, these techniques rely on deriving image-based point clouds and then matching those 3-D points to LiDAR-derived primitives such as points or planes. The reliance of these approaches on image point features is motivated by several fully automated detector, descriptor, and matching techniques such SIFT [21] and SURF [22]. However, the performance of such algorithms significantly degrades when dealing with images with homogenous texture, short baseline, and/or large changes in their viewpoints. It is important to note that these characteristics are dominant for data acquired by the SMART system [5]; thus, current approaches are not applicable for the involved camera/LiDAR integration in this study.

For *in-situ* image/LiDAR integration in indoor environments with homogenous texture, linear features can be considered as a potential primitive. Image line extraction has been well-studied and several reliable and efficient line segment detectors (LSDs) are available [23], [24]. On the other hand, existing region growing [25] or parameter-domain [26] segmentation approaches for extracting LiDAR lines could only perform well when working with high-density point clouds, i.e., they would fail when dealing with sparse point clouds such as those acquired by the SMART system. In addition, due to inherent difficulties in using line primitives such as inaccurate location of endpoints, fragmentation of a single line, and unavailability of a geometric constraint similar to epipolar geometry for point matching [27], establishing conjugate image/LiDAR lines is a challenging task.

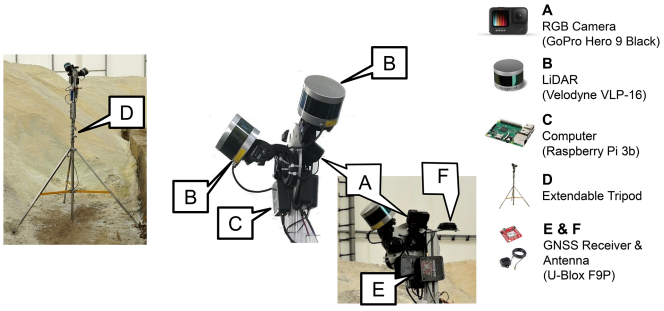


Fig. 1. Integrated SMART system used in this study (system setup for data acquisition within an indoor facility is shown on the left).

This article presents an automated system calibration/registration framework to overcome the limitations of existing approaches through developing strategies for linear features extraction and matching in image and LiDAR data. This framework extracts initial LiDAR lines by detecting edge lines in each acquired point cloud. These linear features are then expanded through region-growing strategies to derive physical lines in the object space. Also, conjugate image/LiDAR lines are identified through a two-step matching strategy that considers the separability among linear features in the image space to derive reliable conjugate lines. Image/LiDAR lines are used in a least squares adjustment (LSA) procedure—i.e., a tool aimed at solving overdetermined systems of equations through minimizing the squared sum of weighted observation errors/residuals—to estimate calibration, registration, and feature parameters.

The rest of this article is structured as follows. The SMART system, data collection procedure, and study sites are introduced in Section II. In Section III, the proposed data processing framework is covered. Experimental results are presented in Section IV. Finally, Section V concludes the article.

II. Smart System and Datasets Description

This section starts with introducing the SMART system used in this study for acquiring image and LiDAR data from indoor stockpiles. Then, system operation and datasets description are covered.

A. Smart System Description

The SMART system, shown in Fig. 1, consists of two Velodyne VLP-16 Puck LiDAR units and one GoPro Hero 9 RGB camera, which are mounted on an extendable tripod pole with a maximum height of 6 m. The VLP-16 Puck scanner consists of 16 radially oriented laser rangefinders that are aligned vertically from -15° to $+15^\circ$ and designed for 360° internal rotation. Two LiDAR sensors with cross orientation are used to increase the area covered by the SMART system from a given data

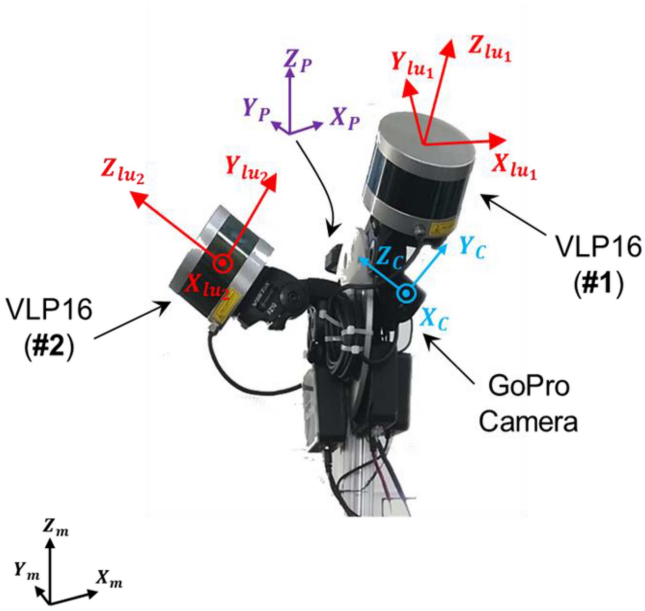


Fig. 2. Illustration of the involved coordinate systems in the SMART system.

acquisition instance. More details regarding the SMART system design and specifications can be found in [5]. More details regarding the SMART system design and specifications can be found in [5].

The SMART system was initially calibrated, where camera IOP together with mounting parameters relating the different sensors were estimated, as described in [5]. Camera IOP including principal distance (c), principal point coordinates (x_p, y_p), and radial and decentering lens distortions (K_1, K_2, P_1, P_2) were derived through an indoor calibration procedure using a test field with several checkerboard targets with known distances. In order to derive the relative position/orientation of the onboard sensors with respect to a common reference frame, a pole coordinate system was first defined, as depicted in Fig. 2. The pole coordinate system together with the mounting parameters of the first LiDAR unit relative to such coordinate system is arbitrarily defined to produce roughly leveled point clouds. The mounting parameters between the remaining sensors and the pole coordinate system were derived by minimizing the discrepancy between conjugate image points and LiDAR planes that were manually extracted from the imagery and LiDAR data, respectively.

B. Smart System Operation

At each instance of data collection, hereafter referred to as a scan, the SMART system captures a pair of LiDAR point clouds along with one RGB image. Fig. 3 shows a sample point cloud and an image captured in a scan. As can be seen in this figure, only a portion of the stockpile is covered by the point cloud. In order to assure as complete coverage as possible of the stockpile facility, multiple scans are required. To do so, the pole is manually or mechanically rotated six times (resulting in seven scans) around its vertical axis in approximately 30° increments, as illustrated in Fig. 4. Depending on the facility

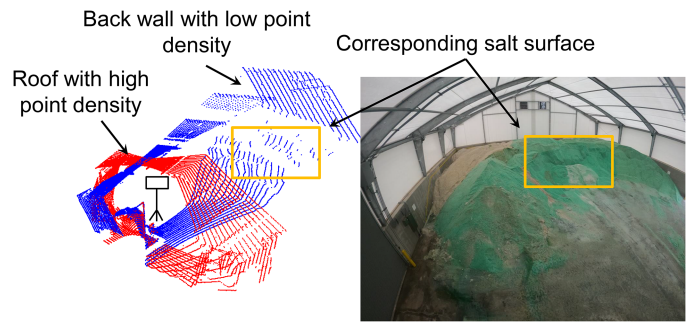


Fig. 3. Sample point cloud (colored by LiDAR unit) and image captured by the SMART system at a given scan.

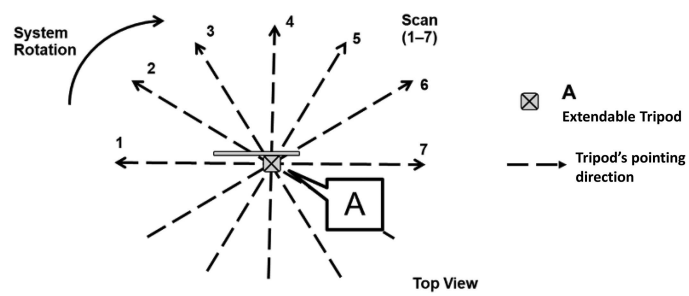


Fig. 4. Data collection procedure for the SMART system (seven scans are acquired at each station).

TABLE I
DATA COLLECTION SITES, FACILITY SIZE, AND NUMBER OF STATIONS/SCANS

Salt Storage Facility	Size (W[1] × L[2] × H[3]) (m)	#[4] of stations	# of scans per station
Rensselaer unit	30 × 79 × 10	3	7
Lebanon unit	26 × 48 × 10.5	2	7
US-231 unit	30.5 × 25.5 × 10	1	7

1: Width, 2: Length, 3: Height, 4: Number

size and shape complexity of the stockpile in question, data collection might be conducted at one or more locations (also referred to as stations). The final 3-D point cloud from the stockpile facility is derived by registering all acquired scans to a common reference frame. It is important to note that such a simple and cost-effective design and data acquisition procedure in the SMART system leads to sparse LiDAR scans, significant variation in the point density, and insufficient overlap between successive scans. These characteristics of the SMART point clouds are considered in the data processing steps.

C. Datasets Description

In this study, three indoor salt storage facilities—managed by the Indiana Department of Transportation—with stockpiles of varying sizes and shapes are used to evaluate the performance of the proposed image/LiDAR integration framework. The locations of the study sites along with a sample aerial view of each facility are illustrated in Fig. 5. The three datasets, denoted as "*Rensselaer*," "*Lebanon*," and "*US-231*" units, are, respectively, located in Rensselaer, Lebanon, and West Lafayette, Indiana, USA. Facility size, number of stations, and number of scans per station for the three datasets are reported in Table I.

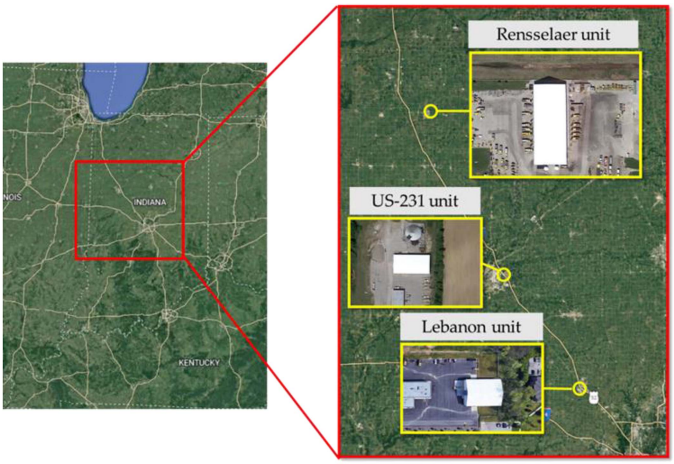


Fig. 5. Locations of the three salt storage facilities used in this study along with a sample aerial view of each facility (aerial photograph adopted from Google Earth Image).

III. Methodology

In this section, the proposed image/LiDAR integration approach for generating colorized point clouds for monitoring indoor stockpiles is presented. The well-aligned colorized point cloud is derived through three main processing blocks, which are illustrated in Fig. 6. The first block in this figure aims at registering acquired point clouds from all scans/stations relative to a common reference frame using planar features as primitives. At this stage, due to the fact that a limited number of planar features exist in each scan, the registered point cloud is susceptible to some misalignments. The second block in Fig. 6 corresponds to the proposed integration approach for the simultaneous refinement of system calibration and registration parameters through establishing conjugate image/LiDAR linear features. Lastly, in the third block, using the refined parameters, a point cloud colorization approach is implemented to assign RGB information to the point cloud. More details about each step of the proposed framework are discussed in the following sections. It is worth noting that the main concept of the approaches pertaining to the first block in Fig. 6 was introduced and discussed in [5]. However, to ease the understanding of the entire workflow, we will start by briefly describing the initial registration step.

A. Initial Point Cloud Registration

The conceptual basis of this step is to use conjugate planar features among all scans at different stations to derive a registered point cloud with a reasonable alignment. This procedure is implemented in four steps, as shown in the first block in Fig. 6. In the first step, the objective is to reconstruct the LiDAR point clouds from S scans at a given station (e.g., $S = 7$) in a common frame, e.g., the coordinate system defined by the pole at the first scan. This is conducted by estimating the pole rotation matrices $R_{p(k)}^{p(1)}$ with k ranging from 2 to 7, using an image-based LiDAR coarse registration approach [5]. More specifically, using a set of conjugate points established between successive images, incremental camera rotation matrices—i.e., $R_{c(k)}^{c(k-1)}$ where $k > 1$—are derived. Then, the pole rotation matrices—i.e., $R_{p(k)}^{p(1)}$—are estimated by considering the boresight angles relating the camera coordinate system and pole body frame denoted as R_c^p.

It is worth mentioning that establishing conjugate points among images acquired in salt storage facilities is quite challenging due to presence of repetitive patterns in the facility. In order to overcome such an issue, a new image matching algorithm—denoted as rotation-constrained matching—was developed in our previous study [5]. In order to reduce the matching ambiguity, this approach exploits the nominal pole rotation to predict the location of a conjugate point in an image for a selected point in another one. Fig. 7 shows two successive scans from the single station in the US-2313 dataset before coarse registration, identified conjugate points (derived from the rotation-constrained image matching) between captured images in these scans, and image-assisted coarse registration result. One should note that the rotation-constrained image matching algorithm is only designed for the SMART system with the assumption that there is no translation between successive scans at a given station. Consequently, derived matched points, although accurate, cannot be used for 3-D reconstruction, and thus, volume estimation of stockpiles.

Next, planar features are extracted from individual scans through a modification of the region-growing segmentation technique proposed by Habib and Lin [28]. In the modified version of the segmentation strategy, points in the LiDAR data are represented as spherical coordinates rather than Cartesian ones. The former provides more meaningful neighborhoods that are robust to point density variations caused by sparse scanning. Given the image-based coarse registration parameters and extracted planar features, conjugate features among scans at a given station can be identified using the similarity of surface orientation and spatial proximity between such planes. Sample plane extraction results for the single station in the *US-231* dataset are shown in Fig. 8.

At this stage, LiDAR scans from the same station are coarsely aligned and planar features are derived from those scans. Next, an automated approach is adapted from the linear feature-based registration approach proposed by Al-Durgham and Habib [29] for the alignment of scans from multiple stations, if available. Multistation registration is sequentially established for two stations at a time, referred to as reference and source stations. Linear features in this step are derived by intersecting nonparallel planes in each station. These lines might or might not correspond to physical lines in the study site. Conjugate linear features between the two stations are then identified and used for the estimation of registration parameters. The reason for exploiting lines instead of direct use of planar features is that the estimation of the registration parameters requires establishing a minimum of three nonparallel plane pairs, whereas only two sets of nonparallel conjugate lines are enough to derive these parameters. Therefore, searching for two conjugate linear features in combined scans at two stations is much easier than searching for three

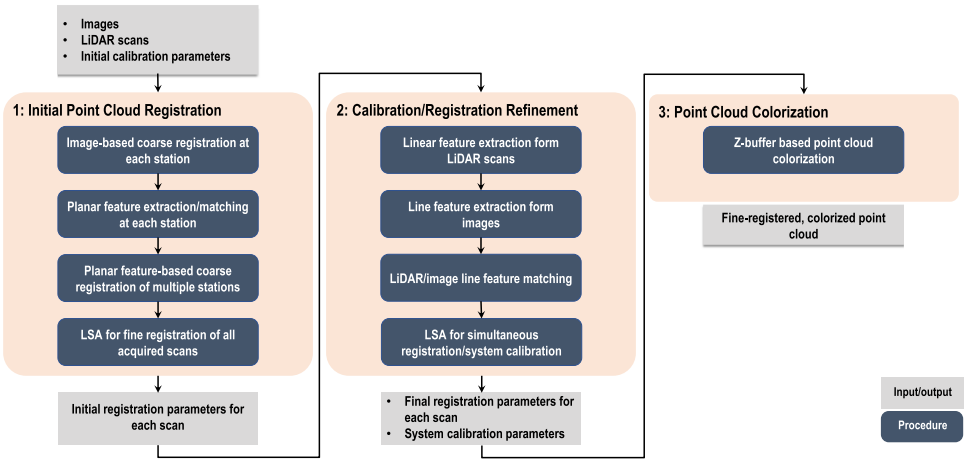


Fig. 6. Proposed data processing framework for deriving well-aligned colorized point clouds from the SMART system.

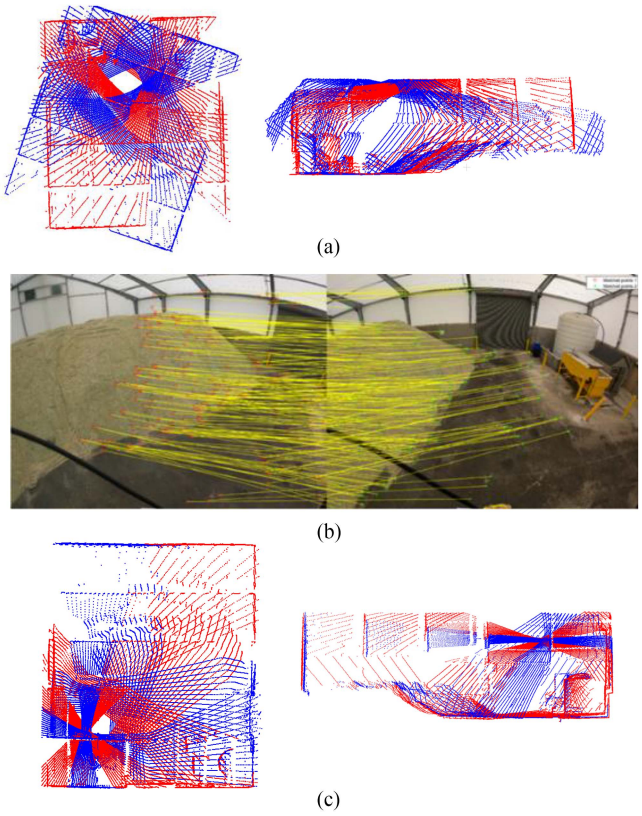


Fig. 7. Image-aided coarse registration results for two successive scans in the *US-231* dataset. (a) Top view (left) and side view (right) of two scans before registration. (b) Identified conjugate points between captured images in these scans. (c) Top view (left) and side view (right) of the two scans after coarse registration (point clouds are colored by scan ID).

conjugate planar features. More details regarding how to establish conjugate lines and use them for the estimation of registration parameters can be found in [29] and [30], respectively.

It is important to note that the generated point cloud in this step might not be well-aligned for two reasons: 1) scans from individual stations are coarsely aligned using the image-aided registration approach; and 2) deriving linear features from planar

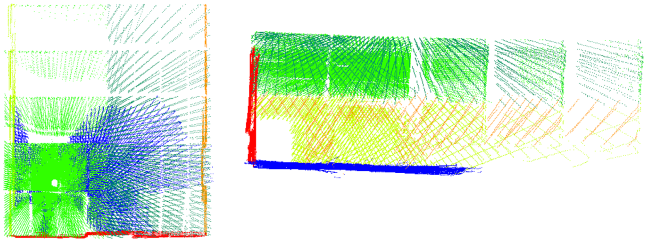


Fig. 8. Top view (left) and side view (right) of extracted/matched planes for the single station in the *US-231* dataset (points are colored by plane ID).

ones leads to the propagation of errors. Fig. 9(a) shows a sample of multistation registration results for two stations in the *Rensselaer* dataset where one can observe coarse alignment among scans. In order to derive a fine-registered point cloud, planar feature matching is first conducted to establish corresponding features among all scans/stations. Finally, a planar feature-based LSA is adopted for simultaneous fine registration of all scans at all stations while solving for the parameters of the best-fitting planes using the approach proposed by Lin et al. [31]. The fine registration results for the scans at the two stations in the *Rensselaer* dataset are shown Fig. 9(b).

B. Calibration/Registration Parameter Refinement

This step of the proposed framework aims at refining the calibration/registration parameters using conjugate linear features among LiDAR scans and imagery. One should note that only a limited number of linear features that were previously derived from plane-to-plane intersection represent physical lines. Thus, only a few of these linear features are visible in the imagery. This step focuses on developing strategies for extracting/matching physical lines in image/LiDAR data using these features for the refinement of calibration/registration parameters, as shown in the middle block of Fig. 6.

Linear feature extraction from LiDAR scans: Linear features in indoor environments are often located in a close proximity to planar features, e.g., exposed ceiling beams in storage facilities.

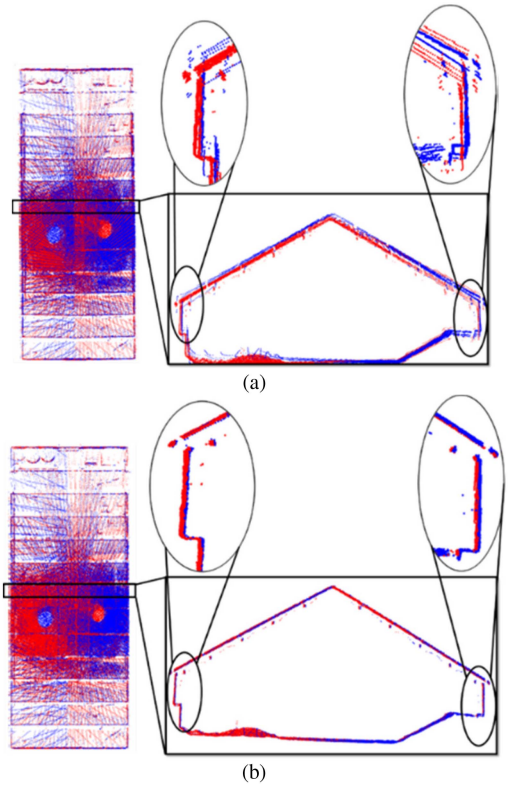


Fig. 9. Registration results for two stations in the *Rensselaer* dataset. (a) Multistation coarse registration and (b) planar feature-based fine registration (point clouds are colored by station ID).

Fig. 10 shows examples of those features with the help of sample images from the three datasets used in this study. When working with sparse point clouds, these linear features cannot be reliably extracted using general segmentation techniques. To overcome the arising challenges from having sparse scans in the SMART system, this study develops a region-growing-based approach that relies on edge points extracted from individual scans captured by each LiDAR unit to derive linear features. This approach consists of three steps, as shown in Fig. 11, which are discussed in this section.

Step 1 aims at generating seed linear segments in the point cloud. To do so, a strategy inspired by the LiDAR odometry and mapping (LOAM) method [32] is implemented. In LOAM, edge and planar points are extracted from individual *LiDAR frames*, i.e., captured data by a multibeam spinning LiDAR sensor at each instance of data collection (a scan is comprised of two LiDAR *frames* for the SMART system). Feature extraction is conducted by calculating the curvature at a point using its local region; points with high curvature value are considered as edge features and those with low value are considered planar points.

In this study, LOAM edge points are used to extract a set of reliable linear features that are extended across the scanlines in LiDAR *frames*. In the implemented approach, given an edge point on a scanline s ($1 \leq s \leq 16$ for the LiDAR units used in the SMART system), a line segment is initiated by finding its closest edge point on scanline $s+1$, provided the two edge points have a similar timestamp (all laser beams rotate simultaneously). Similarly, edge points from adjacent scanlines are sequentially added to the line segment if they have a small timestamp difference to the last point on the segment and their normal distance to the edge line is smaller than a predefined threshold. This threshold is selected as 0.10 m according to the thickness of the utilized linear features (e.g., exposed ceiling beams) in this study. For a line segment to be considered as a valid feature, it should have a minimum of four points, i.e., extracted line segments consist of 4–16 edge points. Fig. 12 shows detected LOAM edge points as well as extracted linear features from a *frame* captured by the first LiDAR unit of the SMART system in the *Rensselaer* dataset. As shown, although not all points on the edge lines are extracted, the majority of the derived linear features are correctly segmented. It should be noted that the LOAM approach might identify points before/after gaps caused by occlusion as edge points, as depicted in the zoomed-in region shown in Fig. 12(a).

In *Step 2* (see Fig. 11), conjugate edge lines derived from different *frames* captured at single or multiple stations are established. To identify collinear segments, initial calibration and registration parameters are used to derive all edge lines in the mapping frame. Fig. 13(a) and (b) shows all extracted edge points and lines, respectively, for the *Rensselaer* dataset where points are colored by their respective station ID. As shown in these figures, only a small portion of edge points on existing lines are segmented as linear features. Therefore, in order to derive full edge lines, extracted linear features from LiDAR *frames* are considered as seed segments and a region growing is conducted on the entire edge points. In this procedure, given a seed segment, edge points with a small normal distance to the linear feature are successively added to the line segment. The point-to-line distance threshold is determined according to the expected accuracy of initial registration/calibration parameters (e.g., 0.10 m in this study). One should note that through the region-growing process, conjugate linear features among scans are implicitly established. Fig. 13(c) depicts the full edge lines resulting from the region-growing procedure on edge lines/points shown in Fig. 13(a) and (b) (points are colored by feature ID). As shown, the majority of existing line segments in the facility have been extracted after the region-growing step. Nevertheless, due to the sparsity of the LiDAR data, a single edge line in the object space might be fragmented during the region-growing process.

Finally, in *Step 3*, a line grouping strategy is implemented to merge linear features that are believed to originate from a single edge line in objects space. In this regard, for each line segment l_i that consists of a set of n points $\mathcal{P}_{l_i} = \{p_1^{l_i}, p_2^{l_i}, \ldots, p_n^{l_i}\}$, line orientation ($\theta_{l_i}$) and length ($L_{l_i}$) are calculated using endpoints $p_1^{l_i}$ and $p_n^{l_i}$. In general, longer lines can be considered to be more reliable than shorter ones. As a result, the grouping process starts with sorting the linear features in a descending order according to their length. This will result in a set of LiDAR lines $\mathcal{L} = \{l_1, l_2, \ldots, l_m\}$, with l_1 and l_m representing the longest and shortest lines, respectively. In the grouping process, for a selected line l_i, we look for its collinear segments in $\mathcal{L}$ by first filtering out lines that show a large angular deviation from l_i (e.g., $\geq 3°$). Such threshold is selected according to the characteristics of the physical lines, e.g., their thickness

Fig. 10. Illustration of sample linear features in indoor facilities highlighted with red ellipses where the majority of line segments are located near planar features.

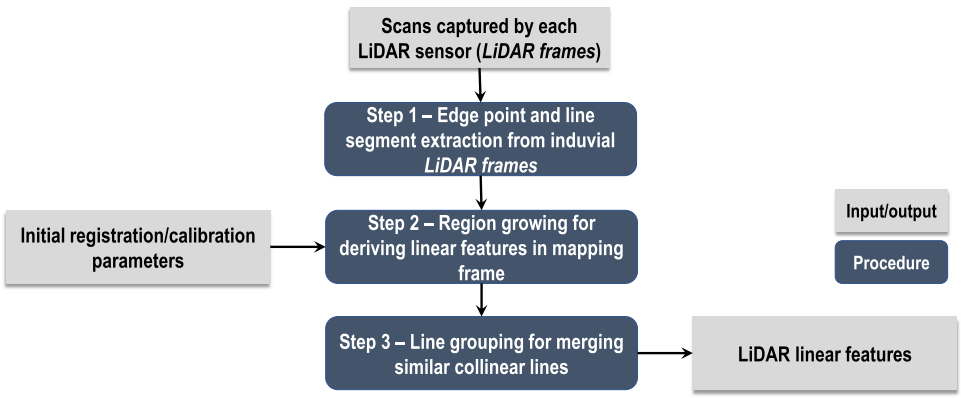


Fig. 11. Workflow of the proposed LiDAR linear feature extraction approach.

and/or straightness. A line l_j that survived the angular proximity check is deemed collinear with l_i ($L_{l_i} \geq L_{l_j}$) if: 1) the root mean square (rms) of the normal distances between the points on l_{l_j}—i.e., $\mathcal{P}_{l_j} = \{p_1^{l_j}, p_2^{l_j}, \ldots\}$—to line l_i is smaller than a threshold (e.g., 0.10 m), and 2) the minimum distance between neighboring endpoints of the two lines is smaller than a predefined threshold. For a more reliable grouping, this threshold is selected as L_{l_j} so that the algorithm does not allow merging of two segments that are far from each other. If l_j is considered to be collinear with l_i, all points in $\mathcal{P}_{l_j}$ are added to $\mathcal{P}_{l_i}$ and line l_j is removed from $\mathcal{L}$. The line grouping procedure is conducted for each line in $\mathcal{L}$ and repeated untilx no more grouping occurs.

In order to remove potential outlier line segments, e.g., lines extracted from the edge points at the boundary of gaps caused by occlusions, a line filtering step is conducted. A line segment is deemed as an outlier if it has a line fitting error larger than a threshold. According to the thickness of the utilized linear features as well as the accuracy of the initial registration parameters, this threshold is set to 0.10 m. Additionally, due to the nature of available linear features in salt storage facilities, segments that are shorter than 0.1 times the length of the longest segment (which is usually one of the exposed ceiling beams) are selected as outliers. Fig. 14 shows the final LiDAR line segments after applying the line grouping/filtering process on the segments shown in Fig. 13(c). Examining this figure, one can observe that the implemented approach results in a set of LiDAR line segments that represent unique physical lines with a good distribution over the study site.

Extracting linear features from imagery: In order to derive image linear features, the LSD algorithm [23] is used in this study. As mentioned earlier, images captured by the SMART system exhibit large distortions; thus, a straight line in the object space might not appear as a straight line in the imagery. Consequently, line detection algorithms would extract a physical line in the object space as several noncollinear lines in the imagery. This issue is illustrated in Fig. 15, where LSD line detection results on a sample original image in the three datasets are shown. To mitigate this problem, the LSD algorithm is applied on rectified images, i.e., images where the distortion parameters have been corrected for. The distortion parameters are established through the previously mentioned indoor calibration procedure, as discussed in Section II. Although such parameters might change overtime, initial estimates would still significantly reduce the impact of distortions. Following the application of LSD, a line grouping strategy similar to the one discussed for LiDAR lines is implemented to merge similar collinear image lines. In order to assure a reliable merging process—i.e., where only line segments that represent the same physical line are merged—conservative thresholds are used in this process. Lastly, short lines, e.g., lines whose length is smaller than 0.1 times the length of the longest segment, are considered outliers

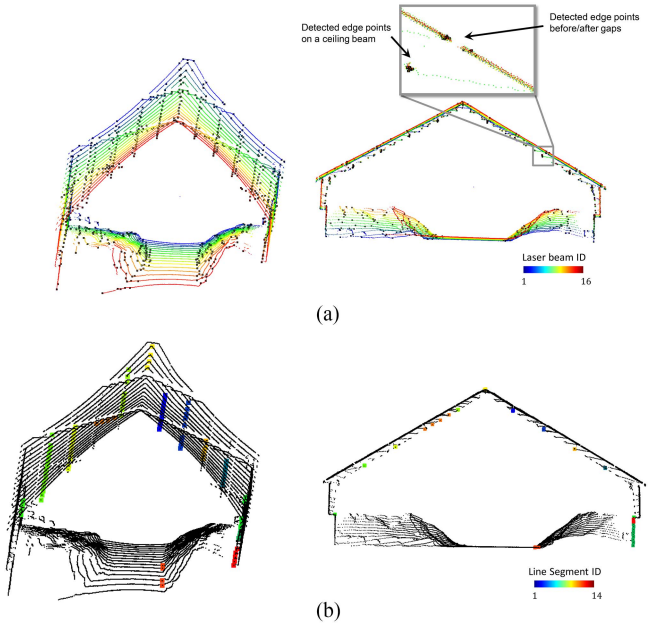


Fig. 12. Perspective view (left) and side view (right) of a *LiDAR frame* in the *Rensselaer* dataset. (a) Points are colored by laser beam ID with LOAM edge points shown in black. (b) Extracted linear features are colored by feature ID and the remaining points are in black.

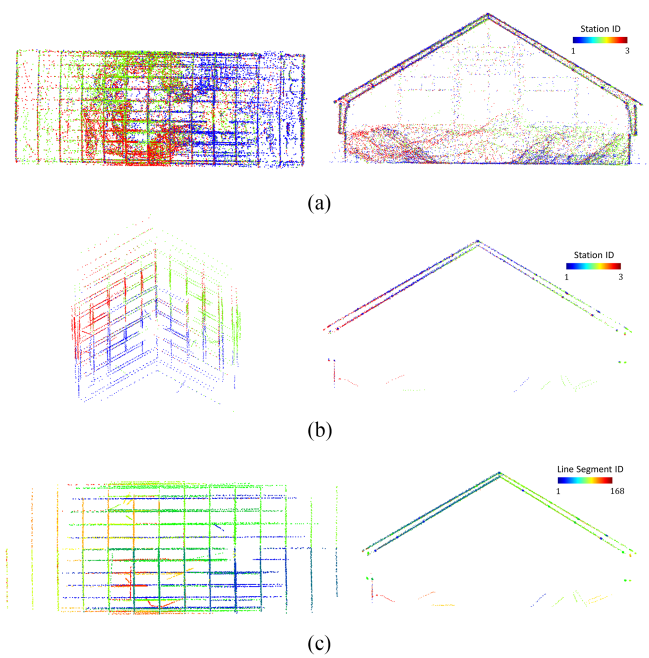


Fig. 13. Top view (left) and side view (right) of edge points/lines extraction results for the entire *Rensselaer* dataset. (a) Extracted edge points colored by station ID, (b) extracted edge lines colored by station ID, and (c) region-growing results (points are colored by linear feature ID).

and removed. Fig. 16 shows the final image lines detected from the images shown in Fig. 15.

LiDAR/image line matching: At this stage, image and LiDAR lines have been extracted. Moreover, LiDAR lines have been matched in the individual LiDAR *frames* from the different stations. This step aims at establishing conjugate linear features

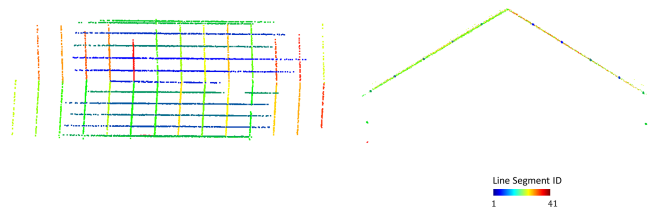


Fig. 14. Top view (left) and side view (right) of final linear features merged/filtered for the *Rensselaer* dataset (points are colored by linear feature ID).

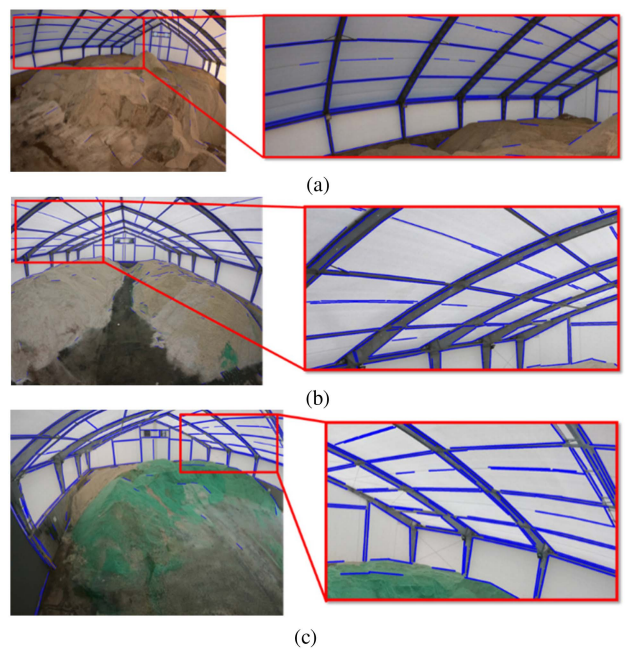


Fig. 15. LSD line detection results (in blue) from sample original images in (a) *Rensselaer* dataset, (b) *Lebanon* dataset, and (c) *US-231* dataset.

between LiDAR data and imagery. The matching process starts by projecting the LiDAR lines onto imagery using the initial calibration/registration parameters. Fig. 17 shows a sample of detected image lines (in blue) and projected LiDAR lines (in red) where, as expected, a misalignment can be observed between the two sets of line segments due to inaccuracy of the initial registration/calibration parameters.

Having LiDAR lines projected to the all images, line matching process is conducted one image at a time. For an image in question, conjugate image/LiDAR lines are established through a two-step strategy. In the first step, for every image line l_i, a potential LiDAR line match is identified. A LiDAR line is considered as a potential match if its projection shows close spatial/angular proximity to the image line in question. Also, there should be no ambiguity in identifying the potential LiDAR line match, i.e., projection of the remaining LiDAR lines should not exhibit relative-close proximity to l_i. In the matching process, among projected LiDAR lines that are in a close angular proximity to the image line in question, the one with the smallest distance to l_i is selected as a potential candidate match l_l (closest projected LiDAR line). In this regard, the distance between the

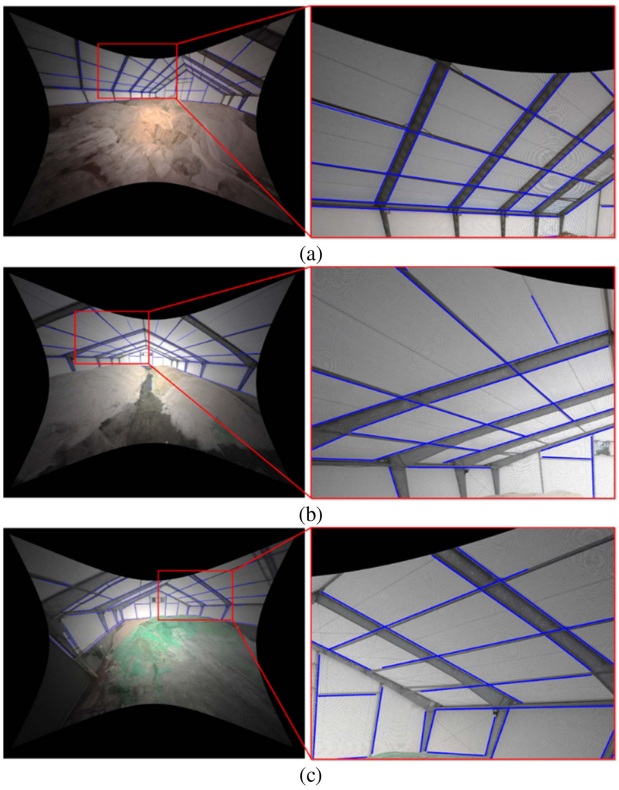


Fig. 16. Final linear features (in blue) detected from sample rectified images in (a) *Rensselaer* dataset, (b) *Lebanon* dataset, and (c) *US-231* dataset.

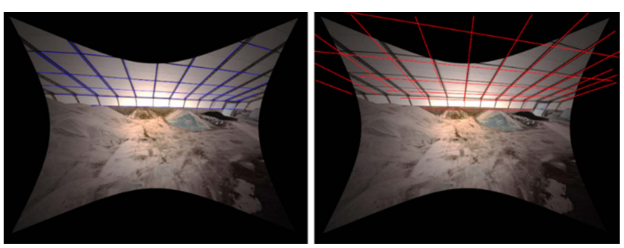


Fig. 17. Extracted image lines in blue (left) and projected LiDAR lines in red (right) for an image in the *Rensselaer* datasets.

two lines is determined as the mean of the normal distances between the endpoints of projected LiDAR line and the image line. Also, the second-closest projected LiDAR line ($l_{l'}$) is the one with the smallest angular/spatial proximity to l_l. Given these three linear features, i.e., image line l_i, closest LiDAR line l_l, and second-closest LiDAR line $l_{l'}$, a matching hypothesis between l_i and l_l is established if the distance between l_i and l_l (d_1) is smaller than a predefined threshold (e.g., $d_1 \leq \rho$ times the image diagonal, where ρ is selected as 0.01 according to the image size while considering the accuracy of the initial calibration/registration parameters), whereas the distance between l_i and $l_{l'}$ (d_2) is significantly larger than d_1 (e.g., $d_2 \geq 2\rho$ times the image diagonal to ensure a good separation between the two projected LiDAR lines). One should note that when multiple line segments with similar orientations are in the same vicinity of the image, all of them would be assigned to a single LiDAR line as a potential match. To ensure reliable matching results, once

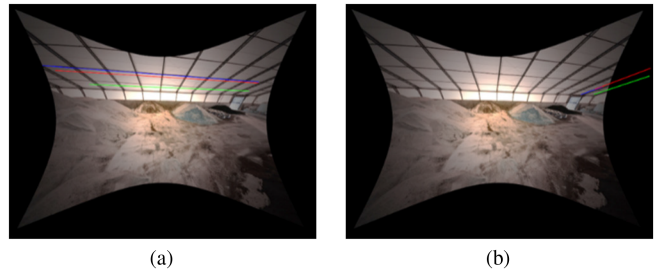


Fig. 18. Two sample image lines (in blue) and their corresponding closest and second-closest projected LiDAR lines in red and green, respectively, for an image in the *Rensselaer* dataset.

TABLE II
MATCHING COST FUNCTIONS AND HYPOTHESES FOR THE IMAGE/LIDAR LINES SHOWN IN FIG. 18

Parameter	Threshold	Fig. 18a	Fig.18b
d_1 **(pixel)**	< 150	99 ✓	37 ✓
d_2 **(pixel)**	> 300	398 ✓	180 ×
Matching hypothesis	N/A	✓	×

candidate LiDAR lines are established for every image line (if possible), in the second step, a potential LiDAR line match is accepted if it is only associated with a single image line within a given image.

The two-step procedure will ensure reliable matches (pairings) between LiDAR/image lines. These reliable matches will correspond to high-resolution regions, i.e., regions with small ground sampling distance (GSD), of the imagery. In other areas, i.e., regions with low-resolution (high GSD), we would have ambiguity in identifying image/LiDAR line matches, i.e., the matching cost function for the second-closest match (d_2) is not significantly different from that for the closest match (d_1). Fig. 18 shows two examples of detected image lines (in blue) along with their corresponding closest and second-closest LiDAR lines (in red and green, respectively). The two matching cost functions for the line segments shown in Fig. 18(a) and (b) are presented in Table II. As reported in this table, image line l_i and LiDAR line l_l in Fig. 18(a) are considered as candidate matches. On the other hand, the matching hypothesis is rejected for the pair in Fig. 18(b) due the small value of d_2, i.e., high ambiguity in identifying image/LiDAR line matches in the low-resolution part of the image.

C. LSA for the Refinement of Calibration and Registration Parameters

Once conjugate linear features are established, an LSA procedure is conducted to refine the system calibration, registration, and feature parameters. In this study, a line in a LiDAR scan is defined by a set of extracted edge points on that line. Among all edge points that belong to a single line (extracted from several scans), the two with the maximum distance in the mapping frame are used to establish the representation scheme of the linear feature in the object space. These two endpoints (line parameters) are denoted as points A and B and are solved for in the LSA procedure. In the meantime, image lines are represented

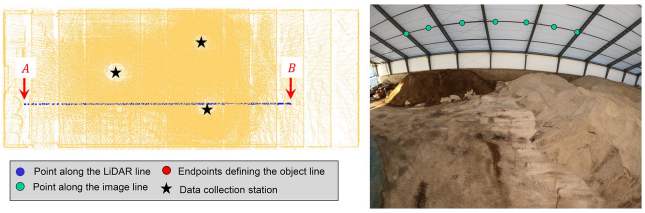


Fig. 19. Set of intermediate points along a conjugate image/LiDAR line in the *Rensselaer* dataset (point cloud is shown in top view).

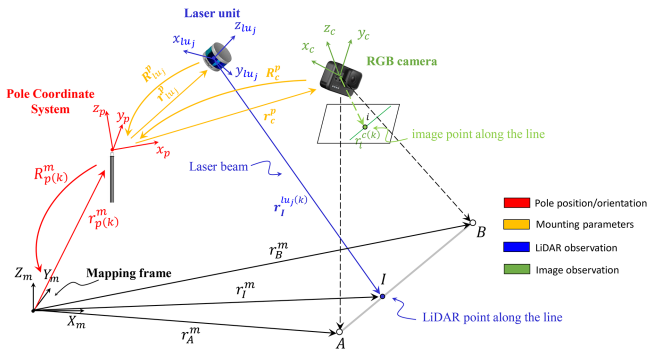


Fig. 20. Schematic illustration of image/LiDAR intermediate points along the linear feature for the SMART system (points I and i represent different intermediate points along the linear feature observed by LiDAR and camera, respectively).

by a sequence of points along the line. Representing an image line by a sequence of points will help in refining lens distortion parameters since the points' deviation from the straight line passing through the endpoints is a function of the distortion. Sample intermediate points along a conjugate image/LiDAR line as well as the representation scheme of this line in the object space (endpoints A and B) are shown Fig. 19.

Two optimization target functions are implemented for minimizing the discrepancies between conjugate image/LiDAR lines while simultaneously solving for the unknowns. The involved quantities in the two target functions are schematically illustrated in Fig. 20. The first optimization target function minimizes the normal distance between the mapping coordinates of the intermediate point I extracted from LiDAR scan to its corresponding linear feature (defined by the endpoints A and B). This constrained is mathematically described in (1), where x represents the L_2-norm of the vector x, r_A^m and r_B^m are the object coordinates of the two endpoints that are determined through the LSA, and r_I^m is the coordinate of LiDAR point I along the line. The LiDAR point coordinates in the mapping frame (r_I^m) are defined using the point positioning equations, as per (2). Considering (1) and (2), one can note that the first LSA target function incorporates the registration parameters, LiDAR mounting parameters, and feature parameters

$$\frac{\|(r_B^m - r_A^m) \times (r_B^m - r_I^m)\|}{\|r_B^m - r_A^m\|} = 0 \tag{1}$$

$$r_I^m = r_{p(k)}^m + R_{p(k)}^m r_{lu_j}^p + R_{p(k)}^m R_{lu_j}^p r_I^{lu_j(k)} \tag{2}$$

where:

$r_{p(k)}^m$ and $R_{p(k)}^m$: position and rotation matrix of the pole body frame relative to the mapping coordinate system (registration parameters) for the kth scan;

$r_{lu_j}^p$ and $R_{lu_j}^p$: lever-arm and boresight rotation matrix relating the jth LiDAR unit coordinate system and pole body frame (j can be either 1 or 2 for the SMART system); and

$r_I^{lu_j(k)}$: position of the intermediate point I with respect to the LiDAR unit frame that is derived from the raw measurements of the jth LiDAR unit at kth scan.

The second optimization target function forces the vector from the perspective center (PC) to an intermediate image point i along the linear feature (r_{PC-i}^m) to lie on the plane defined by the PC and endpoints of the object line (i.e., the plane defined by vectors r_{PC-A}^m and r_{PC-B}^m, as shown in Fig. 20). This constraint is mathematically presented by the triple product in (3), where r_{PC-A}^m is defined by $r_A^m - r_{c(k)}^m$ with $r_{c(k)}^m$ representing the camera position relative to the mapping frame at scan k. Considering the lever-arm relating the camera system to the pole body frame (r_c^p), r_{PC-A}^m is rewritten as $r_A^m - [r_{p(k)}^m + R_{p(k)}^m r_c^p]$ and similarly, r_{PC-B}^m is defined as $r_B^m - [r_{p(k)}^m + R_{p(k)}^m r_c^p]$. Also, r_{PC-i}^m is rewritten as $R_{p(k)}^m R_c^p r_i^{c(k)}$ with R_c^p representing the boresight rotation matrix relating the camera system and pole body frame and $r_i^{c(k)}$ defining the vector from the camera PC to the intermediate image point i along the line in the camera frame captured at scan k

$$\begin{aligned}\left(r_{PC-A}^m \times r_{PC-B}^m\right) \cdot r_{PC-i}^m = \Bigg(&\left[r_A^m - \left[r_{p(k)}^m + R_{p(k)}^m r_c^p\right]\right] \\ &\times \left[r_B^m - \left[{r_{p(k)}}^m + R_{p(k)}^m r_c^p\right]\right]\Bigg) \\ &\cdot R_{p(k)}^m R_c^p \, r_i^{c(k)} = 0\end{aligned} \tag{3}$$

where in $r_i^{c(k)} = \begin{bmatrix} x_i - x_p - dist_{x_i} \\ y_i - y_p - dist_{y_i} \\ -c \end{bmatrix}$, x_p and y_p are the camera's principal point coordinates; c is the camera's principal distance; and $dist_{x_i}$ and $dist_{y_i}$ are distortions in the xy-directions for image point i. Overall, the second minimization target function incorporates the registration parameters, camera IOP, camera mounting parameters, and feature parameters.

It is worth noting that the optimization procedure can be conducted using either LiDAR lines only (for registration purposes, using (1)) or image/LiDAR lines simultaneously [for aligning image and LiDAR data, using (1) and (3)]. In summary, in the LSA procedure, one equation—i.e., (1)—is established for each intermediate point along the LiDAR line. Also, every intermediate point along the image line introduces one equation—i.e., (3)—to the LSA procedure. For each linear feature endpoint, three LiDAR point position equations—i.e., (2)—are also established. The unknown parameters include the camera IOP and mounting parameters, LiDAR mounting parameters, position and orientation of the pole at each scan (except for the reference scan), and ground coordinates of the line endpoints (feature parameters).

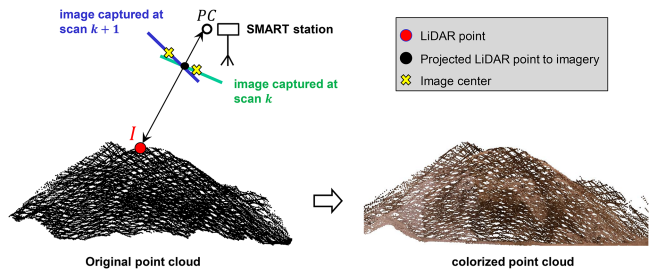


Fig. 21. Schematic illustration of the point cloud colorization process for the SMART system.

D. Point Cloud Colorization

Once the parameter refinement is conducted, a fine-registered point cloud can be generated from all acquired scans at all stations. In the last step of the proposed framework, a colorized point cloud is generated by assigning an RGB value to each point in the final point cloud provided that the point is visible in at least one of the images. A LiDAR point is considered visible in the imagery if it is located in front of the camera and its projection lies within the image format. To colorize a given LiDAR point in the point cloud, its closest station is first determined using the pole position information derived from the LSA. Then, the point in question is projected to all images captured at the closest station. Among projected points, the one with closest distance to the image center, i.e., with less probability of collusion, is used to assign color information to the LiDAR point. This procedure is schematically shown in Fig. 21. In order to avoid the double mapping problem, i.e., problem caused by the fact that a LiDAR point could be occluded in a specific image, the Z-buffer strategy proposed by Ahmar et al. [33] is implemented. In this approach, the point cloud colorization procedure keeps track of projected LiDAR points to a given image location. When several LiDAR points are projected to the same image pixel, the closest point to the PC of the image in question is deemed visible, whereas others are considered occluded. Occluded points are then projected to other images to identify the ones where they might be visible in.

IV. Experimental Results

This section evaluates the capability of the developed image/LiDAR integration strategy for deriving well-aligned colorized point clouds from the SMART system. First, impact of the extracted/matched LiDAR lines (without corresponding image lines) on the registration process is evaluated through a comparison with the results derived from the planar feature-based registration (i.e., initial registration in this study). The capability of the proposed framework for aligning image and LiDAR data is then qualitatively and quantitatively verified. The selected threshold values used in the proposed image/LiDAR integration approach are presented in Table III.

A. LiDAR Fine Registration Results

In this section, only the LiDAR linear features derived from the proposed strategy are used to study their impact on the quality of point cloud alignment compared to that from the planar feature-based LSA. For a quantitative comparison of the two approaches, the following criteria/metrics are reported.

1) *Number of features:* This metric refers to the total number of planar/linear features that are used in the registration process and shows how common these features are within the study site in question.
2) *Minimum number of features per scan:* This criterion refers to the minimum number of features extracted/matched in a LiDAR scan. The criterion represents the ability of the approach in providing sufficient number of features for estimating the registration parameters for all acquired scans.
3) *Matching repeatability:* This metric shows the ratio between the average number of scans among which a given feature is matched and the total number of scans. A high value of this metric is an indication of repeatedly identified conjugate features.
4) *Feature fitting error:* This criterion represents the rms of the normal distances between the points and their respective best-fitted feature derived from the LSA procedure. This value is expected to be within the noise level in the point cloud.

The above-mentioned evaluation criteria for the three datasets are reported in Table IV. As can be seen in this table, a number of extracted linear features are almost three times the number of planar features for all datasets. This observation shows that linear features can be found in large quantities when compared to planar ones for the involved facilities. Also, inspecting the minimum number of features in a scan reported in Table IV, one can note that there are always more than 20 conjugate linear features for every scan of the LiDAR data in all datasets. Since only two nonparallel linear features are enough for deriving the registration parameters of a given scan, such large number of extracted/matched features significantly increases the control information for estimating the unknown parameters in the LSA procedure. On the other hand, the minimum number of planar features in a scan is in the range of 4–6 planes in the all datasets. Having such a small number of planar features in those scans might not always provide adequate control that can lead to

TABLE III
Selected Threshold Values for the Proposed Image/LiDAR Integration Strategy

Threshold description	Threshold value
Normal distance threshold between an edge point to a given edge line for LiDAR line segmentation	0.10 m
Maximum angular deviation threshold for grouping collinear lines (for both LiDAR and image lines)	3°
Maximum spatial separation threshold for grouping collinear lines (for both LiDAR and image lines)	Length of the smaller segment
Line fitting error threshold for removing outlier LiDAR line segments	0.10 m
Minimum line length threshold for removing outlier segments (for both LiDAR and image lines)	0.1 × length of the longest segment
Maximum distance threshold between a given image line to its closest projected LiDAR line (d_1)	0.01×image diagonal
Minimum distance threshold between a given image line to its second-closest projected LiDAR line (d_2)	0.02×image diagonal

TABLE IV
PERFORMANCE COMPARISON BETWEEN PLANAR AND LINEAR FEATURE APPROACHES IN TERMS OF THE NUMBER OF FEATURES USED IN LSA, MINIMUM NUMBER OF FEATURES EXTRACTED/MATCHED IN A SCAN, MATCHING REPEATABILITY, AND FEATURE FITTING ERROR

Dataset (stations/scans)	Approach	Total # of features	Minimum # of features per scan	Matching repeatability	Feature fitting error (m)
Rensselaer (3/21)	Planar-based	13	5	0.59	0.045
	Linear-based	41	26	0.78	0.033
Lebanon (2/14)	Planar-based	9	6	0.65	0.051
	Linear-based	33	22	0.82	0.027
US-231 (1/7)	Planar-based	8	4	0.79	0.071
	Linear-based	38	22	0.69	0.034

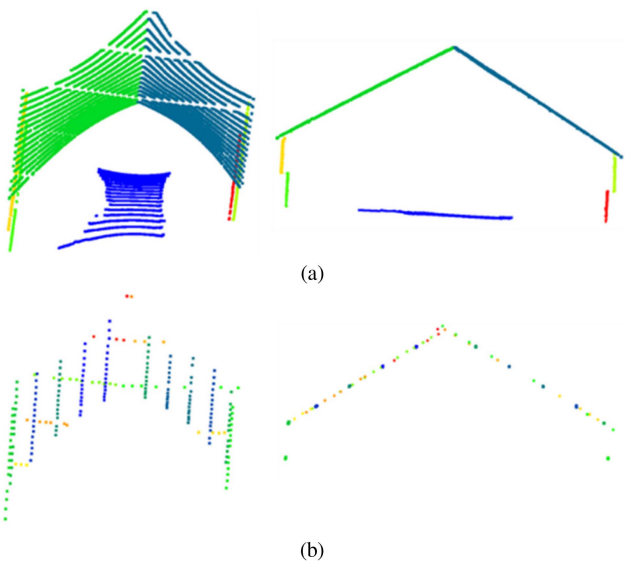


Fig. 22. Perspective view (left) and side view (right) of the extracted/matched features for a LiDAR *frame* in the *Rensselaer* dataset. (a) Seven planar features with a total of ∼20 000 points and (b) 21 linear features with a total of 213 points (points are colored by feature ID).

accurate registration parameters. The number and distribution of planar and linear features used in the LSA procedure are illustrated with the help of Fig. 22 for a *frame* in the *Rensselaer* dataset (previously shown in Fig. 12).

In terms of the matching repeatability, the two approaches show a value in the range of ∼0.6–0.8, indicating that a given feature is matched among the majority (more than half) of the scans. This observation can be attributed to the ability of the SMART system for scanning diverse features in all directions at a given scan (as previously shown in Fig. 3). According to the rms of normal distances presented in Table IV, LSA using planar features shows an rms value within 4.5–7.1 cm, whereas this value is in the range of 2.7–3.4 cm in the case of linear feature registration for all datasets. This level of accuracy is within the noise level of point cloud, i.e., ±3 cm according to the utilized sensor specifications. Such small feature fitting errors indicate the ability of the proposed strategy in extracting accurate LiDAR lines.

Next, the estimated calibration parameters are presented. It is worth noting that to define the pole coordinate system relative to the mapping frame, the mounting parameters of the first LiDAR unit on the SMART system are manually measured and are fixed in the LSA procedure. Table V presents the second LiDAR unit's initial and estimated mounting parameters along with their respective standard deviation (STD) derived from the LSA. The different initial values for the *US-231* dataset compared to the other two are due to using a different setup of the SMART system for this dataset. As one can see in Table V, planar and linear feature approaches result in very similar lever-arm components (within ±2 cm) and boresight angles (within ±0.1°) for all datasets. Also, Table V shows that there are slight improvements in STDs of the estimated parameters when using planar feature registration. This can be explained by the huge redundancy produced by the large number of points on planar features (e.g., ∼20000 points on the planar features compared to 213 points on the linear features as reported/shown in Fig. 22).

RMS of differences between the estimated pole position and orientation at each scan derived from the plane-based and line-based registration approaches pertaining to the three datasets is reported in Table VI. According to these values, we can see that although the positional differences of the registration parameters derived from the two approaches are small (in the range of 1–5 cm), there are large differences between the estimated pole rotation angles, e.g., $\Delta\kappa$ of 0.28° in the US-231 dataset. Since the estimated mounting parameters from the two approaches are very similar (as reported in Table V), with such differences in the registration parameters one would expect notable differences in the alignment quality of the point clouds generated by these approaches. Figs. 23–25 illustrate fine-registered point clouds derived from the planar and linear feature registration approaches for the three datasets. Although these figures show relatively good quality of the generated point clouds from both approaches, a closer inspection reveals that there are some misalignments among the scans in the case of planar feature registration. A better alignment of point clouds when using linear feature approach can be attributed to the large number of lines extracted/matched in the different LiDAR scans as reported in Table IV and/or the possibility of having more outlier points on utilized planar features.

In order to quantitively evaluate the generated point clouds, planar features pertaining to two ceilings, four walls, and one floor (seven in total) are extracted, and a least-squares-based plane fitting is conducted using all extracted points on the planes. Table VII reports the mean rms value of normal distances between points and their respective best-fitted planes. In addition, stockpile volumes are estimated using digital surface models generated from the point clouds as described in [5] and reported in Table VII. According to this table, as expected, a smaller

TABLE V
INITIAL AND ESTIMATED MOUNTING PARAMETERS FOR THE SECOND LIDAR UNIT ON THE SMART SYSTEM THROUGH THE PLANAR AND LINEAR FEATURE APPROACHES

Dataset (station/scan)	Approach	$\Delta X(m)$	$\Delta Y(m)$	$\Delta Z(m)$	$\Delta\omega(°)$	$\Delta\varphi(°)$	$\Delta\kappa(°)$
Rensselaer (3/21)	Initial	−0.16	−0.03	−0.07	−7.10	−57.14	−104.14
	Planar-based	−0.17 ±0.01	−0.04 ±0.01	−0.08 ±0.01	−6.90 ±0.01	−57.04 ±0.01	−104.01 ±0.01
	Linear-based	−0.17 ±0.01	−0.03 ±0.01	−0.08 ±0.02	−6.89 ±0.02	−56.99 ±0.02	−104.02 ±0.02
Lebanon (2/14)	Initial	−0.16	−0.03	−0.07	−7.10	−57.14	−104.14
	Planar-based	−0.16 ±0.01	−0.01 ±0.01	−0.07 ±0.01	−7.13 ±0.01	−57.03 ±0.01	−104.24 ±0.01
	Linear-based	−0.16 ±0.01	−0.03 ±0.01	−0.08 ±0.01	−7.11 ±0.03	−57.01 ±0.03	−104.26 ±0.03
US-231 (1/7)	Initial	−0.19	−0.04	−0.10	−23.40	−57.81	165.42
	Planar-based	−0.20 ±0.01	−0.03 ±0.01	−0.10 ±0.01	−23.14 ±0.01	−57.88 ±0.01	165.59 ±0.01
	Linear-based	−0.18 ±0.01	−0.03 ±0.01	−0.10 ±0.01	−23.37 ±0.06	−57.75 ±0.04	165.47 ±0.06

TABLE VI
RMS OF DIFFERENCES BETWEEN THE ESTIMATED POLE POSITION AND ORIENTATION AT EACH SCAN DERIVED FROM THE PLANE-BASED AND LINE-BASED REGISTRATION APPROACHES

Dataset (station/scan)	RMS δX (m)	RMS δY (m)	RMS δZ (m)	RMS $\delta\omega$ (°)	RMS $\delta\varphi$ (°)	RMS $\delta\kappa$ (°)
Rensselaer (3/21)	0.04	0.02	0.02	0.07	0.24	0.09
Lebanon (2/14)	0.03	0.02	0.02	0.06	0.10	0.12
US-231 (1/7)	0.05	0.03	0.01	0.10	0.12	0.28

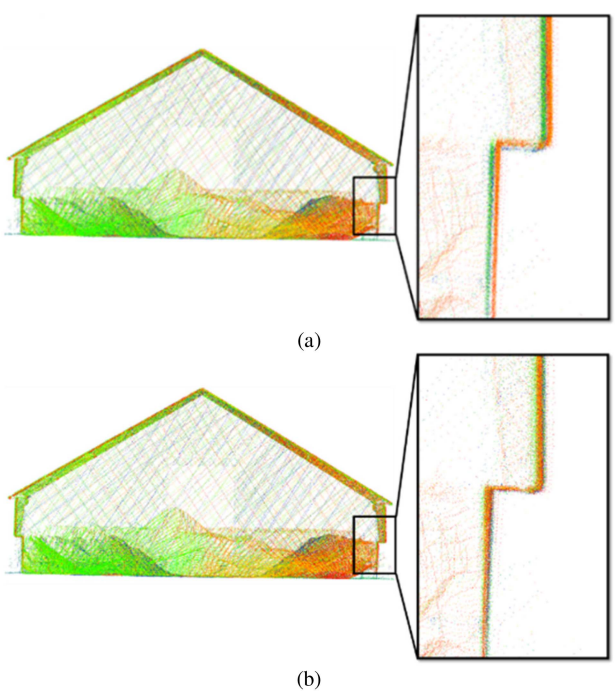


Fig. 23. Point clouds generated for the *Rensselaer* dataset using the estimated parameters from (a) planar and (b) linear features (points are colored by scan ID).

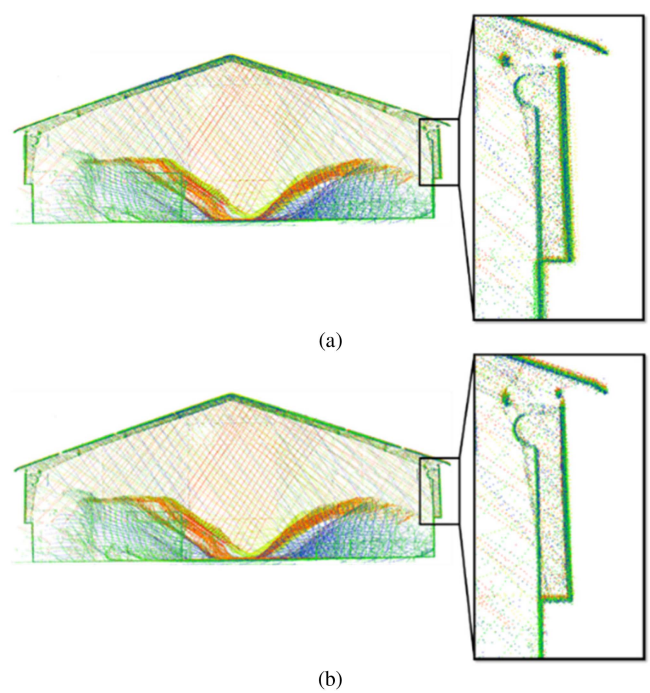


Fig. 24. Point clouds generated for the *Lebanon* dataset using the estimated parameters from (a) planar and (b) linear features (points are colored by scan ID).

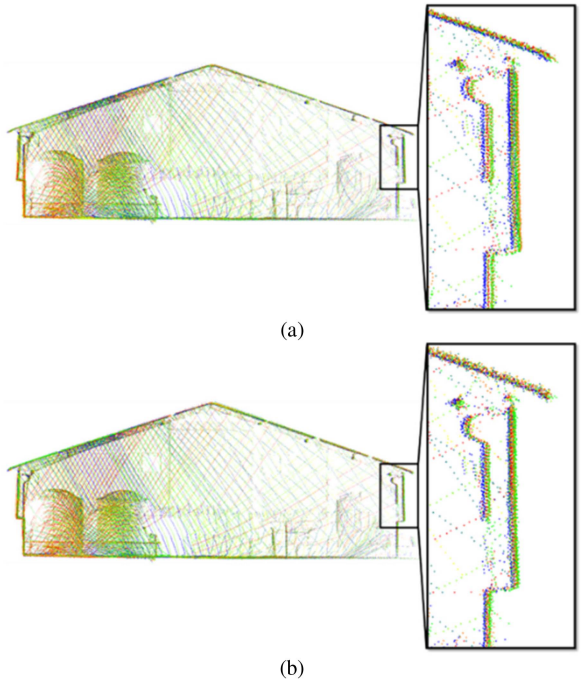


Fig. 25. Point clouds generated for the *US-231* dataset using the estimated parameters from (a) planar and (b) linear features (points are colored by scan ID).

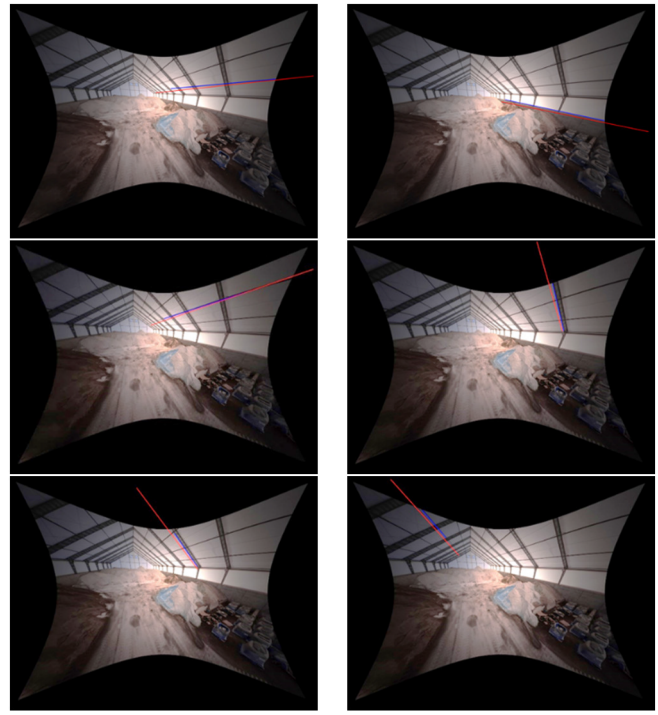


Fig. 26. LiDAR/image line matching results for an image in the *Rensselaer* dataset (image lines are in blue and corresponding LiDAR lines are in red).

TABLE VII
MEAN PLANE FITTING ERROR AND ESTIMATED VOLUMES USING THE TWO APPROACHES

Dataset (stations/scans)	Approach	Mean plane fitting error (m)	Estimated volume (m^3)
Rensselaer (3/21)	Planar-based	0.045	3352.0
	Linear-based	0.035	3350.9
Lebanon (2/14)	Planar-based	0.044	1425.4
	Linear-based	0.039	1424.9
US-231 1/7	Planar-based	0.057	874.5
	Linear-based	0.042	874.6

rms value is observed for the linear feature approach (∼3–4 cm) compared to the planar one for all datasets (∼4–6 cm). However, reported volumes in Table VII show that the two approaches only lead to different volume estimates in the range of 0.1 m^3 (*US-231* dataset) to 1.1 m^3 (*Rensselaer* dataset), which might not be significant for stockpile monitoring applications.

B. Image/LiDAR Integration Results

Having discussed the capability of the proposed approach for identifying sufficient, well-distributed conjugate linear features among all LiDAR scans, this section discusses the image/LiDAR integration results for the three datasets. To do so, the image/LiDAR feature matching results are first presented. Then, the quantitative and qualitative alignment between the point clouds and imagery are analyzed. Fig. 26 shows all identified LiDAR/image line matches (six in total) for a rectified image in the *Rensselaer* dataset. Inspecting this figure, one can observe that all established matches are correct, and they are well-distributed in the image space.

TABLE VIII
IMAGE/LIDAR FEATURE MATCHING RESULTS IN TERMS OF MATCHING RATE AND AVERAGE NUMBER OF CONJUGATE LINES PER IMAGE FOR THE THREE DATASETS

Dataset	Rensselaer	Lebanon	US-231
Image/LiDAR matching rate	83%	79%	55%
Average # of conjugate line per image	6.7	5.9	5.5

The percentage of LiDAR lines for which at least one conjugate image line is identified—i.e., referred to as image/LiDAR matching rate—and the average number of matched LiDAR lines in a given image are reported in Table VIII. The presented image/LiDAR matching rates reveal that nearly 80% of the LiDAR lines were matched with image lines in the *Rensselaer* and *Lebanon* datasets. On other hand, almost half of the LiDAR lines were not matched with image lines in the *US-231* dataset. The lower matching rate for this dataset is caused by the lower overlap between images and LiDAR scans as the data collection is only conducted at one station. Also, considering the average number of matched LiDAR lines reported in Table VIII, one can see that the proposed approach results in an average of ∼5–7 conjugate lines per image for the three datasets. As a result, a large number of constraint equations, i.e., (3), are established in the LSA procedure that contribute to the estimation of involved unknown parameters.

TABLE IX
INITIAL AND ESTIMATED CAMERA DISTORTION PARAMETERS THROUGH THE PROPOSED INTEGRATION APPROACH

Dataset (station/scan)	K_1 $(10^{-8} pixel^{-2})$	K_2 $(10^{-15} pixel^{-4})$	P_1 $(10^{-7} pixel^{-1})$	P_2 $(10^{-7} pixel^{-1})$
Initial	−3.57	−6.95	1.75	−8.90
Rensselaer (3/21)	−3.57±0.03	−7.73±0.03	47.85±0.19	−4.91±0.03
Lebanon (2/14)	−2.48±0.03	−8.35±0.06	9.24±0.27	20.61±0.20
US-231 1/7	−3.11±0.11	−8.19±0.09	45.24±0.21	25.14±0.45

TABLE X
INITIAL AND ESTIMATED CAMERA MOUNTING PARAMETERS THROUGH THE PROPOSED INTEGRATION APPROACH

Dataset (station/scan)	$\Delta X(m)$	$\Delta Y(m)$	$\Delta Z(m)$	$\Delta\omega(°)$	$\Delta\varphi(°)$	$\Delta\kappa(°)$
Initial	0.03	−0.03	0.01	−14.72	−66.45	−103.93
Rensselaer (3/21)	0.07±0.01	−0.02±0.01	0.01±0.01	−14.81±0.02	−64.62±0.01	−103.41±0.02
Lebanon (2/14)	0.01±0.01	−0.08±0.01	−0.04±0.01	−14.02±0.02	−65.48±0.01	−102.21±0.03
US-231 (1/7)	0.02±0.01	−0.05±0.01	−0.01±0.01	−18.53±0.03	−66.04±0.02	−104.52±0.01

It should be noted that the estimated registration as well as LiDAR mounting parameters when using image/LiDAR lines are very similar to those derived from LiDAR lines only (presented in Tables V and VI). Hence, in this section of the experimental results only the estimated camera calibration parameters are presented. In this study, among camera IOP, only distortion parameters are considered as unknowns in the LSA procedure, i.e., principal distance and principal point coordinates are assumed to be stable over time according to what is reported by Zhou et al. [34]. Initial and estimated camera IOP along with their respective STD derived from the linear feature-based LSA are reported in Table IX. The small STD values reported in Table IX indicate precise estimation of the camera distortion parameters for the three datasets. A comparison between the estimated camera IOP among different datasets verifies the hypothesis that these parameters are not stable over time for a consumer-grade camera. Such instability in the parameters is more significant for the decentering lens distortions (P_1 and P_2) compared to the radial distortion parameters (K_1 and K_2).

Table X presents the initial and estimated camera mounting parameters. Similarly, by comparing the estimated mounting parameters among different datasets, one can see that there is a significant change in the estimated parameters, especially in the case of boresight angles. These changes are even more notable for the *US-231* dataset with the largest change in $\Delta\omega$—i.e., $-3.81°$—compared to the derived value from the initial calibration. Such variations in the boresight angles can cause a displacement in the object space with a magnitude of ∼0.66 m at a 10 m distance from the camera (average camera-to-object distance in a typical salt storage facility). The variations in the mounting parameters could be caused by potential movements of the camera due to battery and/or memory card replacement before each data acquisition. Overall, the inconsistency in the camera IOP and mounting parameters necessitates the estimation of these parameters for every dataset acquisition to assure good alignment between LiDAR and image data.

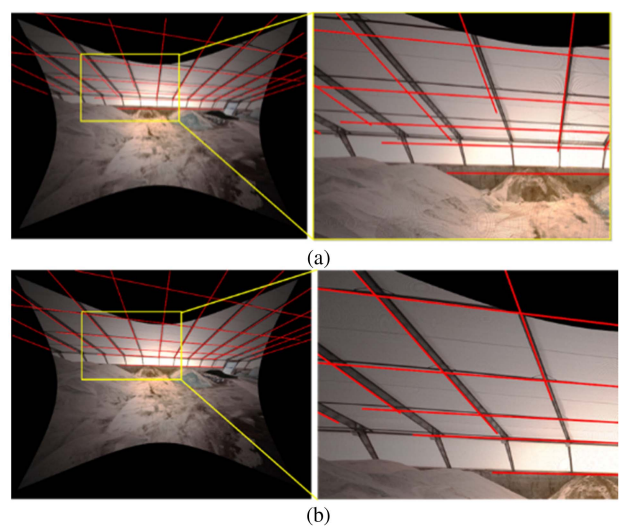
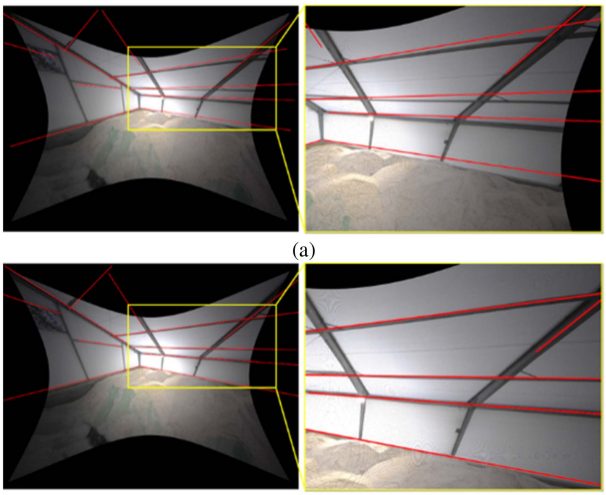


Fig. 27. Projected LiDAR lines for a sample image in the *Rensselaer* dataset: (a) before LSA and (b) after LSA.

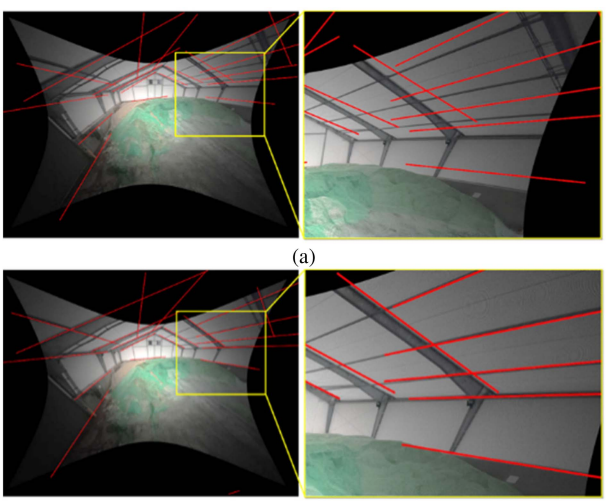
Having discussed the LSA results using image/LiDAR conjugate lines, now the capability of the proposed framework for improving the alignment between corresponding image and LiDAR lines is evaluated. Figs. 27–29 show the projection of LiDAR lines to a sample image before and after LSA for the three datasets. As shown in these figures, when using initial registration/calibration parameters, significant misalignment can be observed between projected LiDAR lines and their respective image lines. Such misalignment is more notable for the *US-231* dataset [see Fig. 29(a)], which supports the previous observation in Table X regarding the significant differences between the



(a)

(b)

Fig. 28. Projected LiDAR lines for a sample image in the *Lebanon* dataset: (a) before LSA and (b) after LSA.



(a)

(b)

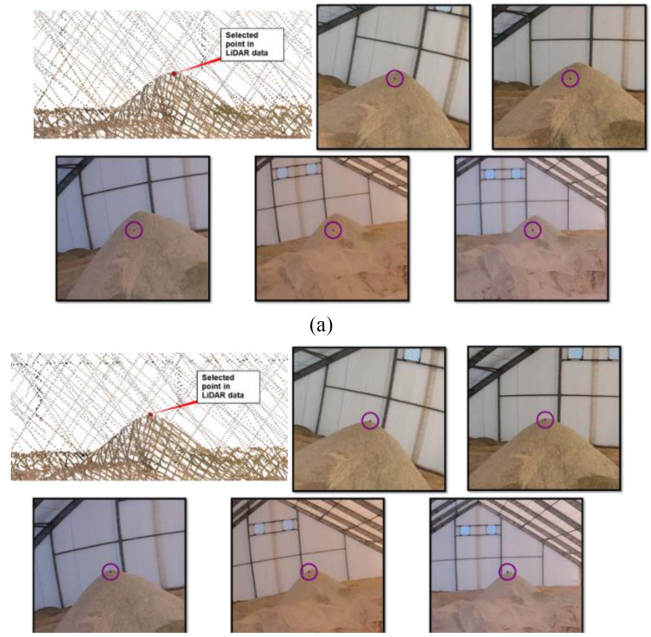
Fig. 29. Projected LiDAR lines for a sample image in the *US-231* dataset: (a) before LSA and (b) after LSA.

TABLE XI
PROJECTION ERROR BETWEEN IMAGE LINES AND THEIR CORRESPONDING LIDAR LINES BEFORE AND AFTER LSA

Dataset	Rensselaer	Lebanon	US-231
Projection error before LSA (pixel)	36.1	31.1	64.4
Projection error after LSA (pixel)	4.4	6.2	5.3

initial and refined boresight angles in this dataset. On the other hand, Figs. 27(b), 28(b), and 29(b) illustrate a good alignment between the projected LiDAR lines and image ones after the LSA procedure.

The rms values of the distances between image lines and their conjugate projected LiDAR lines (denoted as projection error, i.e., d_1 distance discussed in Section III) before and after LSA are presented in Table XI. Considering the reported projection errors before LSA, a significant discrepancy between the image lines and their corresponding LiDAR lines can be observed in all datasets. As expected, such misalignment is more notable for the *US-231* dataset with a magnitude of 64.4 pixels. On the other hand, the projection errors after LSA show that the alignment between the two sets of linear features is improved to the range of 4–7 pixels for all datasets, even for the dataset with only one station (*US-231*).



(a)

(b)

Fig. 30. Sample 2-D analysis for the *Rensselaer* dataset through back-projecting a point in the colorized point cloud to all images where it is visible in (represented by magenta markers): (a) before LSA and (b) after LSA.

Another qualitative alignment check between imagery and LiDAR is assessed through a 2-D analysis by projecting randomly selected points in the LiDAR data to all images where it is visible in Figs. 30–32 illustrate an example of such analysis where the projected points (represented by magenta markers) are derived using the initial registration/calibration parameters [Figs. 30(a)–(32a)] and estimated parameters [Figs. 30(b)–32(b)] for the three datasets. As expected, a misalignment between the projected LiDAR point and its respective image points can be observed when using the initial parameters. On the other hand, using the refined parameters, accurate projections are derived for all selected LiDAR points as shown in Figs. 30(b)–32(b). The projected points in this case are in agreement with their respective image points within ∼5 pixels for all datasets.

Finally, a 3-D qualitative analysis of the image/LiDAR integration results is conducted with a visual inspection of the colorized point clouds. To do so, colorized point clouds using the initial and refined registration/calibration parameters are generated for the three datasets and shown in Figs. 33–35. As expected, inaccuracy in the system parameters adversely affected the quality of the colorized point clouds in all datasets [see the highlighted areas with red ellipses in Figs. 33(a)–35(a)]. On the other hand, Fig. 33(a) and (b) shows good visual quality of the point clouds generated using the refined

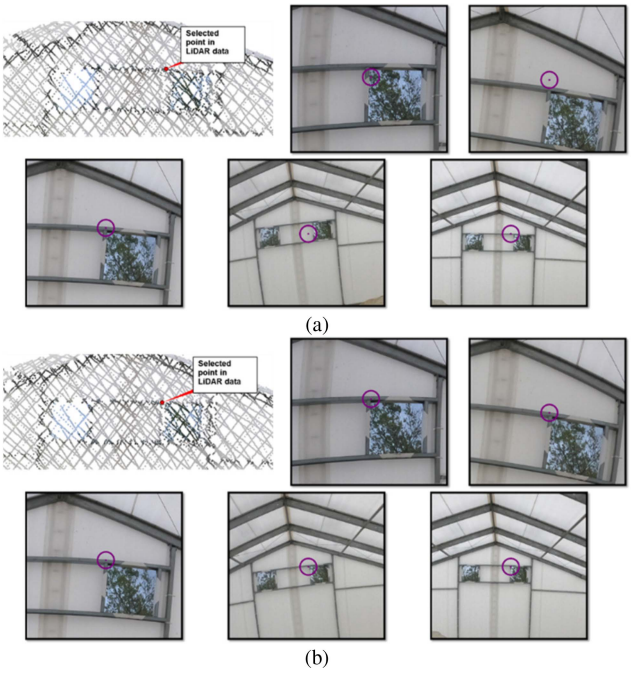


Fig. 31. Sample 2-D analysis for the *Lebanon* dataset through back-projecting a point in the colorized point cloud to all images where it is visible in (represented by magenta markers): (a) before LSA and (b) after LSA.

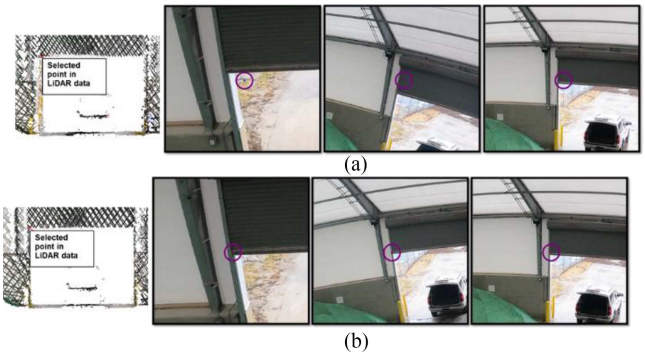


Fig. 32. Sample 2-D analysis for the *US-231* dataset through back-projecting a point in the colorized point cloud to all images where it is visible in (represented by magenta markers): (a) before LSA and (b) after LSA.

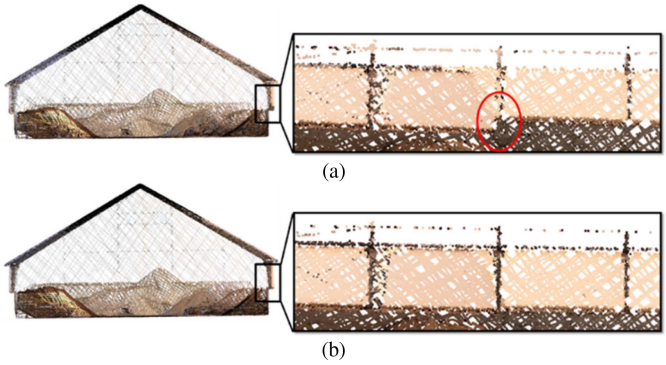


Fig. 33. Side view (left) and front view of a zoomed-in region (right) from colorized point clouds generated for the *Rensselaer* dataset: (a) before LSA and (b) after LSA (red ellipse highlights a falsely colored portion of the point cloud).

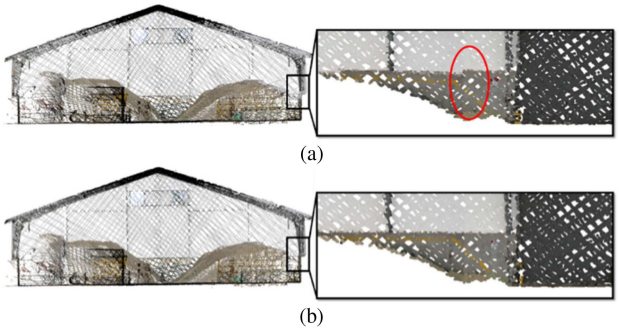


Fig. 34. Side view (left) and front view of a zoomed-in region (right) from colorized point clouds generated for the *Lebanon* dataset: (a) before LSA and (b) after LSA (red ellipse highlights a falsely colored portion of the point cloud).

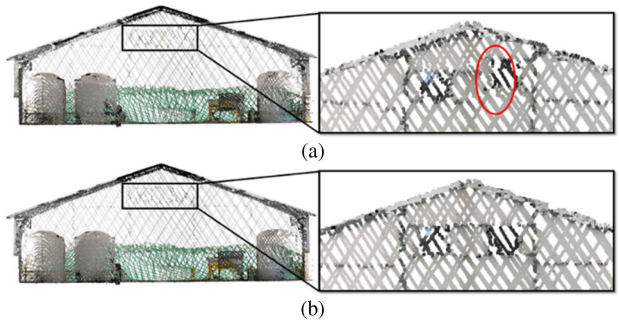


Fig. 35. Side view (left) and a zoomed-in region (right) from colorized point clouds generated for the *US-231* dataset: (a) before LSA and (b) after LSA (red ellipse highlights a falsely colored portion of the point cloud).

parameters. These results suggest that the proposed approach is capable of producing registration and calibration parameters that lead to well-aligned image and LiDAR data from the SMART system.

V. Conclusion and Recommendations for Future Work

In this study, a new image/LiDAR integration using linear features has been proposed for deriving accurate registration and system calibration parameters for a SMART. The key motivation of such development is to produce well-aligned colorized point clouds from acquired data under challenging conditions, e.g., sparse LiDAR point clouds and images with unfavorable geometry (short baseline and/or significant viewpoint change). The proposed strategy is implemented in three main steps: initial registration of LiDAR scans using planar features, calibration/registration parameter refinement using conjugate image/LiDAR lines, and point cloud colorization. The main contributions of this study can be summarized as follows.

1) A new approach for the automated extraction of linear features from sparse LiDAR point clouds has been proposed. The developed strategy extracts seed lines from edge points derived by the LOAM algorithm [32] and exploits them in a region-growing procedure to produce LiDAR lines.

2) A two-step image/LiDAR line matching strategy has been proposed that ensures reliable conjugate features (pairings). These matches correspond to high-resolution regions of the imagery, i.e., they are well-separated from other possible line correspondences.
3) A unified LSA engine has been implemented for system calibration/point cloud registration that can handle either LiDAR scans or integrated image and LiDAR data.

The proposed strategy has been evaluated through experimental results from three datasets with different facility sizes and a number of data acquisition stations. The results showed that using the proposed approach, a large number of linear features with good distribution can be extracted from each LiDAR scan in the SMART system. A comparison between the registration results from linear and planar features revealed that the former can lead to point clouds with better alignments. Nevertheless, both approaches resulted in similar volume estimates of stockpiles for all datasets. In addition, image/LiDAR line matching strategy identified an average of ∼6 conjugate LiDAR lines for each image. LSA-derived registration and system calibration parameters improved the alignment between conjugate image/LiDAR lines from the range of ∼30–65 pixels before LSA, to the range of ∼4–7 pixels. Also, visual inspection of the colorized point clouds generated through the proposed approach showed significant improvements in the quality of the 3-D model compared to that generated using initial calibration/registration parameters.

The proposed approach does not incorporate a line matching outlier removal step. When noncollinear line segments are grouped together and/or initial system calibration parameters are significantly different from their actual values, matching outliers might exist among LiDAR lines and/or image/LiDAR pairings. Those outliers could adversely affect the LSA procedure. To overcome this issue, future work will implement an iterative LSA with a built-in outlier removal strategy. Also, the feasibility of using the proposed approach for facilities with a limited number of planar features (e.g., dome-shaped storage sites [35]) will be investigated.

Acknowledgment

The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein, and do not necessarily reflect the official views or policies of the sponsoring organizations. These contents do not constitute a standard, specification, or regulation.

References

[1] S. M. Hasheminasab, T. Zhou, and A. Habib, "GNSS/INS-assisted structure from motion strategies for UAV-based imagery over mechanized agricultural fields," *Remote Sens.*, vol. 12, no. 3, Feb. 2020, Art. no. 351, doi: 10.3390/rs12030351.

[2] F. He, T. Zhou, W. Xiong, S. M. Hasheminnasab, and A. Habib, "Automated aerial triangulation for UAV-based mapping," *Remote Sens.*, vol. 10, 2018, Art. no. 1952, doi: 10.3390/rs10121952.

[3] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.

[4] H. Hirschmüller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2005, vol. II, pp. 807–814.

[5] R. Manish et al., "Image-aided LiDAR mapping platform and data processing strategy for stockpile volume estimation," *Remote Sens.*, vol. 14, no. 1, 2022, Art. no. 231, doi: 10.3390/rs14010231.

[6] D. D. Lichti, C. Kim, and S. Jamtsho, "An integrated bundle adjustment approach to range camera geometric self-calibration," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 4, pp. 360–368, 2010, doi: 10.1016/j.isprsjprs.2010.04.002.

[7] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2004, vol. 3, pp. 2301–2306.

[8] R. Unnikrishnan and M. Hebert, "Fast extrinsic calibration of a laser rangefinder to a camera," Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR-05-09, 2005.

[9] F. M. Mirzaei, D. G. Kottas, and S. I. Roumeliotis, "3D LIDAR-camera intrinsic and extrinsic calibration: Identifiability and analytical least-squares-based initialization," *Int. J. Robot. Res.*, vol. 31, no. 4, pp. 452–467, 2012, doi: 10.1177/0278364911435689.

[10] S. Verma, J. S. Berrio, S. Worrall, and E. Nebot, "Automatic extrinsic calibration between a camera and a 3D Lidar using 3D point and plane correspondences," in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2019, pp. 3906–3912.

[11] S. Mishra, G. Pandey, and S. Saripalli, "Extrinsic calibration of a 3D-LIDAR and a camera," in *Proc. IEEE Intell. Veh. Symp.*, 2020, pp. 1765–1770.

[12] L. Zhou, "A new minimal solution for the extrinsic calibration of a 2D LIDAR and a camera using three plane-line correspondences," *IEEE Sensors J.*, vol. 14, no. 2, pp. 442–454, Feb. 2014.

[13] C. Guindel, J. Beltran, D. Martin, and F. Garcia, "Automatic extrinsic calibration for lidar-stereo vehicle sensor setups," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, 2017, pp. 1–6, doi: 10.1109/ITSC.2017.8317829.

[14] J. Beltran, C. Guindel, A. de la Escalera, and F. Garcia, "Automatic extrinsic calibration method for LiDAR and camera sensor setups," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 10, pp. 17677–17689, Oct. 2022.

[15] D. Scaramuzza, A. Harati, and R. Siegwart, "Extrinsic self calibration of a camera and a 3D laser range finder from natural scenes," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2007, pp. 4164–4169.

[16] P. Moghadam, M. Bosse, and R. Zlot, "Line-based extrinsic calibration of range and image sensors," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 3685–3691.

[17] S. Schneider, T. Luettel, and H. J. Wuensche, "Odometry-based online extrinsic sensor calibration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2013, pp. 1287–1292.

[18] Z. Taylor and J. Nieto, "Motion-based calibration of multimodal sensor extrinsics and timing offset estimation," *IEEE Trans. Robot.*, vol. 32, no. 5, pp. 1215–1229, Oct. 2016.

[19] J. Li, B. Yang, C. Chen, and A. Habib, "NRLI-UAV: Non-rigid registration of sequential raw laser scans and images for low-cost UAV LiDAR point cloud quality improvement," *ISPRS J. Photogramm. Remote Sens.*, vol. 158, pp. 123–145, 2019, doi: 10.1016/j.isprsjprs.2019.10.009.

[20] Q. Pentek, P. Kennel, T. Allouis, C. Fiorio, and O. Strauss, "A flexible targetless LiDAR–GNSS/INS–camera calibration method for UAV platforms," *ISPRS J. Photogramm. Remote Sens.*, vol. 166, pp. 294–307, Aug. 2020, doi: 10.1016/j.isprsjprs.2020.05.014.

[21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, pp. 91–110, 2004.

[22] H. Bay, T. Tuytelaars, and L. Van Gool, "LNCS 3951 - SURF: Speeded up robust features," in *Proc. 9th Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.

[23] R. G. Von Gioi, J. Jakubowicz, J. M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, Apr. 2010.

[24] C. Akinlar and C. Topal, "EDLines: A real-time line segment detector with a false detection control," *Pattern Recognit. Lett.*, vol. 32, no. 13, pp. 1633–1642, 2011, doi: 10.1016/j.patrec.2011.06.001.

[25] A. Dimitrov and M. Golparvar-Fard, "Segmentation of building point cloud models including detailed architectural/structural features and MEP systems," *Automat. Construction*, vol. 51, pp. 32–45, 2015, doi: 10.1016/j.autcon.2014.12.015.

[26] S. Filin and N. Pfeifer, "Segmentation of airborne laser scanning data using a slope adaptive neighborhood," *ISPRS J. Photogramm. Remote Sens.*, vol. 60, no. 2, pp. 71–80, 2006, doi: 10.1016/j.isprsjprs.2005.10.005.

[27] B. Fan, F. Wu, and Z. Hu, "Robust line matching through line-point invariants," *Pattern Recognit.*, vol. 45, no. 2, pp. 794–805, 2012, doi: 10.1016/j.patcog.2011.08.004.

[28] A. Habib and Y. J. Lin, "Multi-class simultaneous adaptive segmentation and quality control of point cloud data," *Remote Sens.*, vol. 8, no. 2, 2016, Art. no. 104, doi: 10.3390/rs8020104.

[29] K. Al-Durgham and A. Habib, "Association-matrix-based sample consensus approach for automated registration of terrestrial laser scans using linear features," *Photogramm. Eng. Remote Sens.*, vol. 80, no. 11, pp. 1029–1039, 2014, doi: 10.14358/PERS.80.11.1029.

[30] A. Habib, M. Ghanma, M. Morgan, and R. Al-Ruzouq, "Photogrammetric and lidar data registration using linear features," *Photogramm. Eng. Remote Sens.*, vol. 71, no. 6, pp. 699–707, 2005, doi: 10.14358/PERS.71.6.699.

[31] Y. C. Lin et al., "Processing strategy and comparative performance of different mobile lidar system grades for bridge monitoring: A case study," *Sensors*, vol. 21, no. 22, 2021, Art. no. 7550, doi: 10.3390/s21227550.

[32] J. Zhang and S. Singh, "LOAM: Lidar odometry and mapping in real-time," in *Proc. Robot., Sci. Syst. Conf.*, 2014, pp. 1–9, doi: 10.15607/RSS.2014.X.007.

[33] F. Ahmar, J. Jansa, and C. Ries, "The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM," *Int. Arch. Photogramm. Remote Sens.*, vol. 32, pp. 16–22, 1998.

[34] T. Zhou, S. M. Hasheminasab, R. Ravi, and A. Habib, "LiDAR-aided interior orientation parameters refinement strategy for consumer-grade cameras onboard UAV remote sensing systems," *Remote Sens.*, vol. 12, 2020, Art. no. 2268, doi: 10.3390/rs12142268.

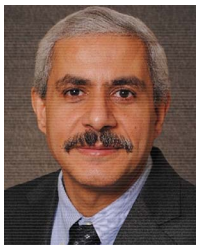
[35] J. Liu, S. M. Hasheminasab, T. Zhou, R. Manish, and A. Habib, "An image-aided sparse point cloud registration strategy for managing stockpiles in dome storage facilities," *Remote Sens.*, vol. 15, 2023, Art. no. 504, doi: 10.3390/rs15020504.

Seyyed Meghdad Hasheminasab received the Ph.D. degree in geomatics engineering from the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA, in 2022.

He is currently a Computer Vision Engineer with the OPPO US Research Center, Palo Alto, CA, USA. His research interests include the fields of structure from motion and simultaneous localization and mapping for applications in precision agriculture, infrastructure monitoring, and augmented/virtual reality.

Tian Zhou received the B.S. degree in remote sensing from Wuhan University, Wuhan, China, in 2015, and the M.S. degree in civil engineering in 2019 from Purdue University, West Lafayette, IN, USA, where he is currently working toward the Ph.D. degree in geomatics with the Lyles School of Civil Engineering.

His research interests include sensor calibration, camera/LiDAR integration, SLAM, image/LiDAR-based 3-D reconstruction using mobile mapping systems, LiDAR point cloud processing, and hyperspectral orthorectification.



Ayman Habib received the M.Sc. and Ph.D. degrees in photogrammetry from The Ohio State University, Columbus, OH, USA, in 1993 and 1994, respectively.

He is currently the Thomas Page Professor with the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA. His research interests include the fields of terrestrial and aerial mobile mapping systems using photogrammetric and LiDAR remote sensing modalities, UAV-based 3-D mapping, and integration of multimodal, multiplatform, and multitemporal remote sensing data for applications in transportation, infrastructure monitoring, environmental protection, precision agriculture, digital forestry, resource management, and archeology.