

## MULTI-PRIMITIVE TRIANGULATION OF AIRBORNE AND TERRESTRIAL MOBILE MAPPING IMAGE AND LIDAR DATA

T. Zhou<sup>1</sup>, J. Liu<sup>1</sup>, S. Shin<sup>1</sup>, A. Habib<sup>1\*</sup>

<sup>1</sup>Lyles School of Civil Engineering, Purdue University, West Lafayette, USA - (zhou732, liu2845, shin337, ahajib)@purdue.edu

**KEY WORDS:** Mobile Mapping Systems, Uncrewed Aerial Vehicles, Camera LiDAR Integration, System Calibration, Trajectory Enhancement.

### ABSTRACT:

Over the last few decades, mobile mapping systems (MMS) such as uncrewed aerial vehicles (UAVs) and terrestrial platforms have been demonstrated in collecting geospatial data for a wide range of applications. MMS continue to evolve due to the unprecedented developments in sensor technology and emerging application domains. Integration of image and LiDAR data acquired by these systems can provide a comprehensive 3D model of the area of interest. However, ensuring good alignment of derived products from single or multiple platforms is crucial. Although many studies have been conducted in this area, there is still a need for a comprehensive integration approach that minimizes discrepancies between imagery and LiDAR data due to inaccurate calibration parameters or trajectory artifacts. To address this issue, a tightly-coupled camera/LiDAR integration workflow denoted as Unified Multi-Sensor Advanced Triangulation (UMSAT) is proposed. UMSAT can handle point, linear, and areal features derived from imaging and ranging systems while utilizing the position and orientation information provided by the navigation unit. This paper explores the feasibility of the proposed framework in two applications – archaeological mapping and geometric documentation of transportation corridors – for improving the quality of derived data/products from imaging and ranging remote sensing systems. Experimental results demonstrate that the UMSAT framework successfully aligns multi-temporal, multi-sensor, and multi-platform geospatial data.

### 1. INTRODUCTION

Remote sensing technologies are increasingly being utilized for a variety of applications. With the emergence of both passive and active remote sensing modalities, a diverse range of useful information can be derived. However, spaceborne and airborne remote sensing platforms do not provide reasonable spatial/temporal resolution at an affordable cost. Therefore, modern mobile mapping systems (MMS) have emerged as promising platforms due to developments in sensor technology and increasing application fields that could benefit from collected geospatial data. For example, the continuous developments in direct geo-referencing technology (i.e., integrated Global Navigation Satellite Systems and Inertial Navigation Systems – GNSS/INS) and remote sensing modalities (i.e., passive and active imaging sensors in the visible and infrared range) are providing the professional geospatial community with opportunities to generate accurate 3D information with a rich set of attributes. These advances are also coupled with improvements in the sensors' performance, reduction in the associated cost, and miniaturization of such sensors. We are also enjoying the emergence of promising platforms such as uncrewed aerial vehicles (UAVs) and terrestrial MMS; providing near-proximal and proximal, respectively, sensing capabilities. UAVs are becoming popular for small-area mapping due to their low cost, ease of deployment, high maneuverability, and advancements in georeferencing technologies and imaging/ranging sensors. Wheeled platforms also offer their unique benefits, such as long operation time, flexibility, small sensor-to-object distance, and reduced operational risk, making them a suitable alternative for certain applications. These advantages have facilitated the use of such platforms in various applications such as geometric documentation of transportation corridors, smart agriculture, coastal monitoring, digital forestry, transportation management, infrastructure monitoring, and non-destructive archaeology.

MMS usually carry imaging/ranging sensors like RGB, multispectral/hyperspectral cameras, and LiDAR units. RGB cameras, which are based on a frame imaging mechanism, capture images in three color channels and provide spectral information in a two-dimensional raster data structure. In contrast, multispectral/hyperspectral cameras might use push-broom technology to capture 1D images with fine frequency bands across the spectrum. To derive 3D information, RGB

cameras are often used in combination with Structure from Motion (SfM) (Westoby et al., 2012) and dense matching algorithms (Furukawa & Ponce, 2009) to generate dense point clouds. However, their performance is dependent on adequate overlap/side-lap among neighboring images and establishing sufficient matches. Multispectral/hyperspectral cameras, on the other hand, do not provide overlap between captured 1D images, making it difficult to derive 3D information. To address this limitation for imaging sensors, LiDAR sensors, which directly provide 3D points with high geometric accuracy but lack spectral/color information, are often integrated with RGB cameras to create a more complete 3D representation of the object space. Incorporating a digital camera and a LiDAR unit onboard UAVs and wheeled platforms provides more capabilities, which enhances the processes of feature extraction, scene understanding, and visualization of derived products. For instance, the fusion of camera and LiDAR data was adopted for archaeological mapping in several studies (Kadhim & Abed, 2021; Vilbig et al., 2020). Camera/LiDAR integration was also used for transportation corridor documentation for road detection, traffic sign detection/recognition, and automated sign retro-reflectivity condition evaluation (Ai & Tsai, 2016; Caltagirone et al., 2019; L. Zhou & Deng, 2014).

For successful integration of multi-temporal data/products from various sources, the positional quality be carefully addressed. This requires accurate calibration of the system, including the sensor's interior orientation parameters (IOP) and mounting parameters that relate the sensors to the INS' Inertial Measurement Unit (IMU) body frame, as well as trajectory information. For image/LiDAR integration, several studies have focused on system calibration techniques. These techniques can be classified into target-based and target-less approaches, depending on the utilized features. Target-based approaches utilize custom-built targets, such as planar checkerboards (Verma et al., 2019), while target-less approaches rely on identifying conjugate natural points and linear features in the scene (Moghadam et al., 2013). Fully-automated camera/LiDAR calibration frameworks have also been developed, such as motion-based approaches that utilize visual odometry (Schneider et al., 2013) or SfM (Glira et al., 2016; Zhou et al., 2021) to establish conjugate features between image and LiDAR data. These techniques might also refine trajectory information to achieve the best alignment between imagery and LiDAR data.

\* Corresponding author

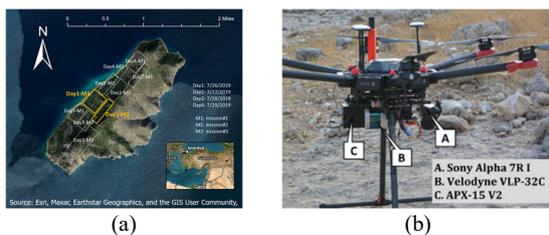
Although considerable research has been conducted on integrating camera and LiDAR data, there is still a need for an in-situ, comprehensive approach that can reduce the discrepancies caused by inaccurate system calibration parameters and trajectory artifacts. In response, this study proposes a tightly-coupled camera/LiDAR integration workflow for GNSS/INS-assisted UAV and wheeled remote sensing systems. The proposed approach – Unified Multi-Sensor Advanced Triangulation (UMSAT) – is designed to handle point, linear, and areal features from both imaging (such as frame cameras and push-broom scanners) and ranging modalities. By utilizing the position and orientation information provided by the GNSS/INS unit, UMSAT can accurately integrate multi-temporal data/products from different modalities. In this study, real datasets from two application domains – archaeological mapping and transportation corridor documentation – are utilized to validate the feasibility of the proposed strategy.

## 2. ACQUISITION SYSTEMS AND DATASETS DESCRIPTION

To evaluate the performance of UMSAT, a complex archaeological site was surveyed using camera and LiDAR units mounted on a UAV system. For the geometric documentation of transportation corridors, imagery/LiDAR data were acquired along a highway by UAV and wheeled systems. The following subsections provide details about the study sites, specifications of the used MMS, and acquired datasets.

### 2.1 Dana Island Datasets

Dana Island, which was part of ancient Rough Cilicia in southern Turkey, is characterized by steep slopes, dense vegetation, as well as the presence of numerous cisterns and burial rooms. The site is covered by ten flight missions (as shown in Figure 1a) for cistern detection purposes. An in-house developed UAV system was used for the data acquisition. The UAV, as shown in Figure 1b, is equipped with a LiDAR scanner – Velodyne VLP-32C, and a digital camera – Sony  $\alpha$ 7R. The LiDAR unit is mounted on the UAV with its vertical axis parallel to the flight direction. The Sony  $\alpha$ 7R is a 36.4-megapixel (MP) camera. The camera, which is set up on the UAV while having its optical axis pointing in the nadir direction, is triggered at a frame period of 1.5 seconds. The LiDAR and camera units are directly georeferenced by an APX-15 UAV V2 GNSS/INS unit. With an IMU data rate of 200 Hz, the unit provides a post-processing accuracy of 2-5 cm for position,  $0.025^\circ$  for roll/pitch angles, and  $0.080^\circ$  for heading angle under open sky conditions. Due to the isolated location of the island, a local Trimble base station was established for differential GNSS post-processing. However, obvious misalignment among different missions was observed when using the base station. Therefore, two sample datasets (hereafter denoted as *Dana-M1* and *Dana-M2*), highlighted in orange in Figure 1a, will be used in this study to evaluate the performance of UMSAT. The flying height, flying speed, spatial coverage, and collected data of the two missions are listed in Table 1.



**Figure 1.** Illustration of (a) the study site and acquired datasets and (b) the UAV-based MMS used for archaeological mapping.

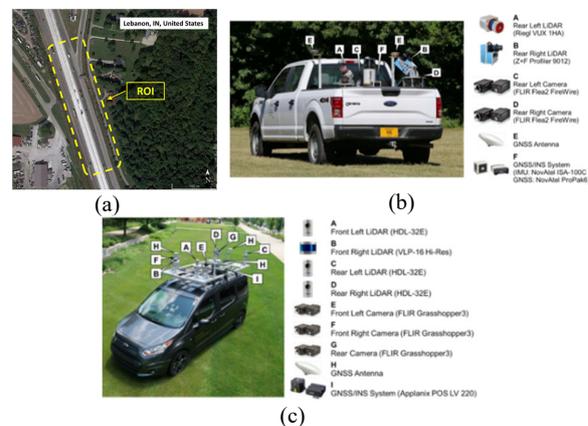
Dataset	Flying height (m)	Flying speed (m/s)	Number of Collected LiDAR Points (in millions)	Number of Captured Images	Spatial coverage (ha)
<i>Dana-M1</i>	45-65	~6.0	~76	514	~6.5
<i>Dana-M2</i>	30-50	~5.8	~87	518	~7.8

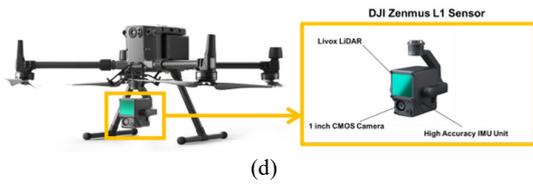
**Table 1.** Specifications of the two datasets acquired by the UAV system.

### 2.2 I-65 Highway Datasets

For the geometric documentation of transportation corridors, a 0.5-mile segment along the I-65 highway close to Lebanon, IN, United States was selected as the study site, as shown in Figure 2a. Two in-house developed wheeled MMS, Purdue wheel-based mobile mapping system-Ultra High Accuracy (PWMMS-UHA) and Purdue wheel-based mobile mapping system-High Accuracy (PWMMS-HA), were involved in this study. The PWMMS-UHA, as displayed in Figure 2b, is equipped with two single-beam LiDAR scanners: Riegl VUX 1HA and Z+F Profiler 9012. Two rear-facing FLIR Flea2 FireWire cameras are installed; they have a maximum image resolution of 5.0 MP and are synchronized to capture images at a frame rate of 0.75 seconds. All sensors are directly georeferenced by a NovAtel ProPak6 and ISA-100C GNSS/INS unit. The GNSS/INS post-processing positional accuracy is 1-2 cm with an attitude accuracy of  $0.003^\circ$  for pitch/roll and  $0.004^\circ$  for heading. The PWMMS-HA, as shown in Figure 2c, includes four multi-beam LiDAR scanners: three Velodyne HDL-32E and one Velodyne VLP-16 Hi-Res. Three FLIR Grasshopper3 9.1 MP cameras are also mounted on the PWMMS-HA: two forward-facing and one rear-facing. The cameras are synchronized to capture one frame per second per camera. The PWMMS-HA sensors are directly georeferenced by an Applanix POS LV 220 GNSS/INS unit. After post-processing, a positional accuracy of 2 cm with an attitude accuracy of  $0.020^\circ$  and  $0.025^\circ$  for the roll/pitch and heading can be achieved.

In addition, an off-the-shelf UAV system – DJI M300 equipped with the Zenmuse L1 LiDAR sensor – is used (Figure 2d). The Zenmuse L1 integrates a Livox LiDAR module, a camera, and an IMU on a 3-axis stabilized gimbal. The UAV camera has a 1-inch CMOS with a 24 mm focal length and a maximum image resolution of 20.7 MP. The IMU unit has a measurement rate of 200 Hz. After post-processing, the unit provides a position accuracy of 1.0-1.5 cm and attitude accuracy of  $0.025^\circ$  and  $0.15^\circ$  for pitch/roll and heading, respectively. Three datasets were acquired using these systems on different dates. Table 2 summarizes the relevant information for the datasets acquired from the PWMMS-UHA, PWMMS-HA, and DJI UAV systems.





**Figure 2.** Illustrations of (a) the study site along the I-65 highway, and used MMS including in-house developed wheeled systems (b) PWMMS-UHA and (c) PWMMS-HA, as well as (d) off-the-shelf DJI M300 UAV equipped with the Zenmuse L1 (adapted from DJI website).

Platform	Date	Driving/flight Speed (mph)	Number of Collected LiDAR Points (in millions)	Number of Captured Images
PWMMS-UHA	2022/12/02	~50	~23	52
PWMMS-HA	2022/07/10	~50	~63	158
DJI UAV*	2022/08/02	~8.5	~85	88

\* Above ground flying height is 70 meters.

**Table 2.** Specifications of acquired datasets for the wheeled and UAV systems.

### 3. METHODOLOGY

Proper georeferencing of the involved sensors, together with comprehensive modeling of the point positioning equations relating their measurements to the respective ground coordinates, is the key to multi-modal geospatial data integration. The point positioning equations for LiDAR and imaging systems are established in two steps. First, the laser beam or imaging ray is defined relative to the sensor coordinate system based on the sensor measurements and IOP. For a LiDAR unit, the former include laser range/pointing direction while the latter denote the parameters describing its encoder mechanism. For a camera, sensor measurements refer to image coordinate measurements and IOP include principal point coordinates, principal distance, and distortion parameters. Second, the position and orientation of the laser beam/imaging ray relative to the mapping frame are established through the Exterior Orientation Parameters (EOP). When a GNSS/INS unit is used, the EOP are derived using the post-processed trajectory and mounting parameters relating these sensors to the corresponding IMU body frame.

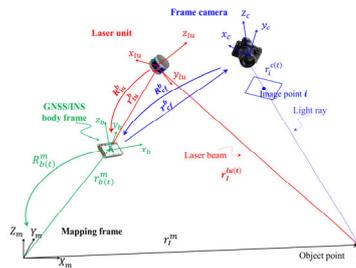
The point positioning models for LiDAR and frame camera units are illustrated in Figure 3. In this figure,  $r_l^{lu(t)}$  denotes the position of the footprint of a laser beam, emitted at time  $t$ , relative to the laser unit frame;  $r_i^{c(t)}$  represents the imaging ray for point  $i$  relative to the camera coordinate systems at time  $t$ ;  $r_{lu}^b/R_{lu}^b$  and  $r_c^b/R_c^b$  represent the lever arm and boresight rotation matrix relating the laser unit/camera and IMU body frame coordinate systems, respectively;  $r_b^m$  and  $R_b^m$  are the position and orientation of the IMU body frame relative to the mapping frame at time  $t$ ; and  $r_l^m$  is the coordinates of object point  $l$  in the mapping frame. The point positioning models are mathematically presented in Equations (1) and (2). The derivation of  $r_l^{lu(t)}$  is based on the range/pointing direction measurements of the LiDAR unit as well as its IOP, while  $r_i^{c(t)}$  is derived from the image coordinates of point  $i$  ( $x_i$  and  $y_i$ ) and camera IOP, including the principal point coordinates of the used camera ( $x_p$  and  $y_p$ ), principal distance ( $f$ ), as well as distortions in the  $x$  and  $y$  coordinates for image point  $i$  ( $dist_{x_i}$  and  $dist_{y_i}$ ). Although the

frame camera is used as an illustration example, the model is applicable to line camera, where the  $y$  coordinate is always constant – e.g.,  $y_i = 0$  for systems with the scan line vertically below the camera perspective center. As mentioned in Section 1, while LiDAR can directly provide 3D information from a single beam, image-based 3D reconstruction involves an unknown scale factor  $\lambda(i, c, t)$ , which needs to be estimated using overlapping imagery.

$$r_l^m = r_b^m + R_b^m r_{lu}^b + R_b^m R_{lu}^b r_l^{lu(t)} \quad (1)$$

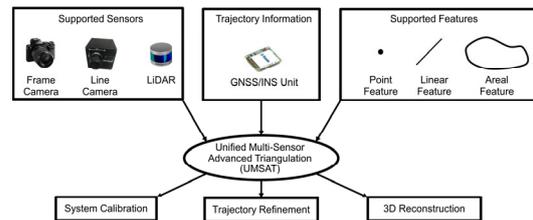
$$r_l^m = r_b^m + R_b^m r_c^b + \lambda(i, c, t) R_b^m R_c^b r_i^{c(t)}, \quad (2)$$

$$r_i^{c(t)} = \begin{bmatrix} x_i - x_p - dist_{x_i} \\ y_i - y_p - dist_{y_i} \\ -f \end{bmatrix}$$



**Figure 3.** Schematic diagram of point positioning principle for LiDAR and frame camera onboard a GNSS/INS-assisted MMS.

The presented point positioning equations suggest that accurate system calibration parameters (including sensor IOP and mounting parameters) and trajectory information are critical for producing well-georeferenced data from LiDAR and imaging systems. To enhance the positional quality of collected data, a system-driven triangulation strategy – unified multi-sensor advanced triangulation (UMSAT) – is proposed. As illustrated in Figure 4, this framework aims to minimize discrepancies among conjugate features, including point, linear, and areal features, captured by different sensor modalities from either single or multiple GNSS/INS-assisted systems. The following subsections will first introduce the feature extraction and matching strategies used in the two sample applications involving imaging and LiDAR sensors onboard UAVs and wheeled vehicles. Then, the optimization framework for system calibration and trajectory enhancement will be presented.



**Figure 4.** Schematic diagram of the functionality of the proposed UMSAT framework.

#### 3.1 Feature Extraction and Matching

The success of camera/LiDAR integration is dependent on whether features are reliably extracted and matched among different modalities and systems. Since we rely on natural features in the study site, feature extraction and matching strategies vary depending on the environment in question. Typically, image-based object points can be established through SfM algorithms for cameras mounted on UAV systems with a

near-nadir view. As for LiDAR data, areal features are the most commonly used primitives. These features will be used for the Dana Island datasets. On the other hand, linear features might be available for some specific applications – e.g., lane markings and light poles in geospatial data covering transportation corridors.

**Feature extraction and matching for the Dana Island datasets:** For each UAV dataset collected at Dana Island, the LiDAR point cloud is first derived using the GNSS/INS-derived trajectory. To extract reliable ground patches, which will be used as areal features, a terrain model generation strategy is used (Shin et al., 2023). The strategy can handle rugged terrain with sudden elevation change, dense vegetation cover, and/or the presence of underground structures while separating above-ground (AG) and bare-earth (BE) point clouds. Ground patches are then obtained from the BE point cloud through an iterative local plane fitting process. Given that multi-temporal LiDAR datasets have reasonable alignment in the planimetric direction, derived planar patches and corresponding inlier LiDAR observations from different datasets are considered conjugate features as long as their normal vectors have similar orientation. For imagery data, a GNSS/INS-assisted SfM algorithm is conducted (Hasheminasab et al., 2020) to derive image-based object points and corresponding image tie points for each dataset. To ensure the alignment between imagery and LiDAR data, correspondences between image-based object points and LiDAR areal features are established for each dataset. Considering that the site has dense vegetation cover, a point-areal feature pair is considered valid only if the planar patch corresponds to exposed BE points, which are identified based on the existence of AG points in a local neighborhood. The ratio of exposed/unexposed BE points in a local neighborhood is then used to determine exposed planar patches. Finally, for a given image-based object point, its closest planar patch is identified. A valid point-areal feature pair is established if the planar patch is exposed and the normal distance from the object point to the plane is smaller than a predefined threshold. The point-areal matching process is schematically illustrated in Figure 5. The extracted image-based object points, LiDAR planar patches, and correspondences between them from the two UAV datasets are used for the subsequent UMSAT processing.

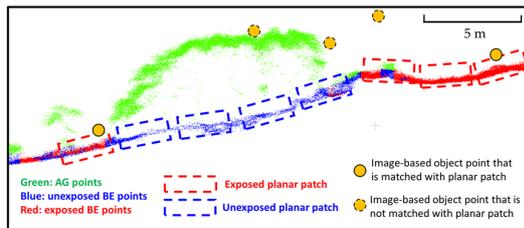


Figure 5. Illustration of identifying corresponding image-based object points and LiDAR planar patches.

**Feature extraction and matching for the I-65 highway datasets:** For datasets acquired from UAV and wheel-based MMS over I-65 Highway, lane markings in the site are considered as linear features. Specifically, skip-lines are modeled as individual linear features. Edge lines, on the other hand, are divided into short straight-line segments, as shown in Figure 6. First, a geometry-based strategy (Cheng et al., 2020) is used to extract lane marking from the LiDAR data of each system. Then, considering that the initial alignment among different datasets is reasonable, lane markings derived from the PWMMS-UHA, PWMMS-HA, and DJI LiDAR data are matched according to their spatial proximity. For imagery data from wheeled systems, corresponding linear features are manually extracted through the

back-projection of endpoints of LiDAR linear features onto the image space. Lane marking features mainly provide control information on the vertical and across driving directions as they are approximately parallel to each other. To have control along the driving direction, four light poles from LiDAR data are manually extracted and treated as linear features. For the imagery data from DJI UAV, in addition to the manually extracted linear features, image tie points are derived with the corresponding object points through an SfM strategy.

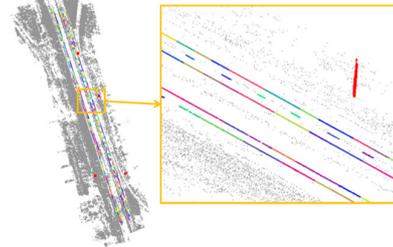


Figure 6. Established lane markings (randomly colored by the feature ID) and four poles (colored in red) from the DJI LiDAR data, as well as the image-based object points (colored in gray) derived from the SfM processing of DJI imagery.

### 3.2 UMSAT Optimization Framework

This subsection starts by introducing the cost functions related to point, linear, and areal features. As mentioned before, imaging systems mainly utilize point features established from SfM algorithms. In UMSAT, the back-projection error serves as the cost function for an image-based object point  $I$  and its corresponding conjugate image points. Back-projection error is derived from the point positioning equation by reformulating it into Equation (3), followed by eliminating the unknown scale factor, as presented in Equation (4). Based on this equation, the back-projection errors – i.e., the differences between the observed image coordinates and predicted ones using estimated unknowns – are minimized in the least squares adjustment (LSA) process.

$$r_i^{c(t)} = \frac{1}{\lambda(i, c, t)} \left[ R_b^c R_m^{b(t)} (r_l^m - r_{b(t)}^m - R_{b(t)}^m r_c^b) \right] = \frac{1}{\lambda(i, c, t)} \begin{bmatrix} N_x \\ N_y \\ D \end{bmatrix} \quad (3)$$

$$x_i = -c \frac{N_x}{D} + x_p + dist_{x_i} \quad (4)$$

$$y_i = -c \frac{N_y}{D} + y_p + dist_{y_i}$$

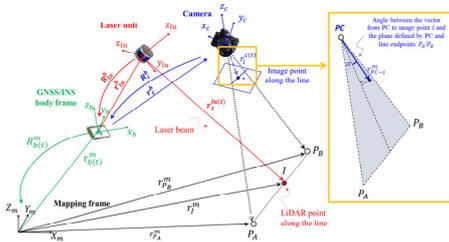
Linear features, which can be established from both imagery and LiDAR data, are represented by two endpoints  $P_A$  and  $P_B$ . Figure 7 provides a schematic illustration of the optimization target functions related to conjugate LiDAR and image observations, along with the relevant quantities involved in these models. The first type of target function seeks to minimize the normal distance between the mapping coordinates of a LiDAR point  $I$  and the respective linear parametric model defined by endpoints  $P_A$  and  $P_B$ . This constraint is mathematically described in Equation (5), with  $\|x\|$  denoting the  $L_2$ -norm of the vector  $x$ ,  $r_{P_A}^m$  and  $r_{P_B}^m$  representing the object coordinates of the two endpoints. On the other hand, the image target function describes the fact that the vector from the perspective center (PC) to an intermediate image point  $i$  along the linear feature ( $r_{PC-i}^m$ ) lies on the plane defined by the PC and endpoints of the object line (i.e., the plane defined by vectors  $r_{PC-P_A}^m$  and  $r_{PC-P_B}^m$ , as shown in Figure 7). This constraint is expressed mathematically in Equation (6), using the triple product of the above three vectors. In this equation,  $r_{PC-P_A}^m$

is defined by  $r_{P_A}^m - r_{c(t)}^m$ , where  $r_{c(t)}^m$  represents the camera position relative to the mapping frame at time  $t$  and is derived from trajectory information and camera mounting parameters; while  $r_{P_C-i}^m$  is represented as  $R_{b(t)}^m R_c^b r_i^{c(t)}$ . To evaluate the residual of this constraint following the LSA, the angle  $\alpha$  between the vector  $r_{P_C-i}^m$  and the plane defined by the PC and object line endpoints  $P_A/P_B$  (as shown in Figure 7) is computed.

$$\frac{\|(r_{P_B}^m - r_{P_A}^m) \times (r_{P_C-i}^m - r_{P_A}^m)\|}{\|r_{P_B}^m - r_{P_A}^m\|} = 0 \quad (5)$$

$$(r_{P_C-A}^m \times r_{P_C-B}^m) \cdot r_{P_C-i}^m = \left( [r_{P_A}^m - [r_{b(t)}^m + R_{b(t)}^m r_c^b]] \right) \times [r_{P_B}^m - [r_{b(t)}^m + R_{b(t)}^m r_c^b]] \cdot R_{b(t)}^m R_c^b r_i^{c(t)} = 0 \quad (6)$$

$$AX_I + BY_I + CZ_I + D = 0 \quad (7)$$



**Figure 7.** Schematic illustration of image/LiDAR points along the linear feature for the SMART system (points  $I$  and  $i$  represent the points along the linear feature observed by LiDAR and camera, respectively).

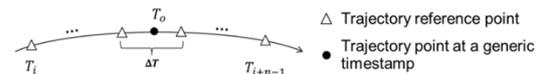
Typically, areal features are solely used for LiDAR sensors as they do not provide redundant information for an imagery observation. In UMSAT, areal features are represented as planes, and the target function is designed to minimize the normal distance between the LiDAR point  $I$  to the corresponding areal feature, as mathematically defined in Equation (7), with the plane parameters denoted as  $(A, B, C, D)$  and  $(X_I, Y_I, Z_I)$  referring to the coordinates of LiDAR point  $I$  in the mapping frame,  $r_I^m$ .

In addition to minimizing the residuals from the camera or LiDAR observation to its respective feature primitives, UMSAT has the capability to ensure the consistency of corresponding features in the object space. Specifically, taking the Dana Island datasets as an example, although no conjugate features could be derived from the two modalities, image-based object point  $P$  and LiDAR-based planar patch correspondences were established. The alignment of imagery and LiDAR data can be guaranteed by minimizing the normal distance between the image-based object point  $P$  and the respective parametric model of the areal feature. This constraint is also applicable to endpoints of a linear feature extracted from imagery data, assuming it rests on an areal feature defined by LiDAR points.

The abovementioned optimization functions for camera and LiDAR observations involve the respective sensor's system calibration parameters, trajectory information at the time of observation, and parameters defining the object point/linear/areal feature. In the UMSAT, trajectory information is refined by estimating corrections ( $\delta r_{b(t)}^m / \delta R_{b(t)}^m$ ) to the original position and orientation parameters obtained from GNSS/INS post-processing. However, it is not recommended to solve for trajectory corrections for every camera/LiDAR observation, as it may lead to over-parametrization in the LSA. Instead, since the platform has a relatively smooth trajectory with moderate

dynamics, the original high-frequency trajectory is down-sampled using a time interval  $\Delta T$ , resulting in trajectory reference points shown in Figure 8. The corrections to the trajectory parameters at a specific observation timestamp are then modeled as a  $p^{th}$ -order polynomial function of the unknown corrections at its  $n$  neighboring trajectory reference points, as expressed in Equation (8). Here, the trajectory corrections at a generic timestamp,  $T_0$ , denoted as  $\delta \theta_{b(T_0)}^m$ , are a function of the timestamps and trajectory corrections of its  $n$  neighboring trajectory reference points. The down-sampling time interval, polynomial order, and number of neighboring trajectory reference points are selected based on the characteristics of the platform dynamics.

$$\delta \theta_{b(T_0)}^m = f(T_0, T_i, \dots, T_{i+n-1}, \delta \theta_{b(T_i)}^m, \dots, \delta \theta_{b(T_{i+n-1})}^m) \quad (8)$$



**Figure 8.** Down-sampled trajectory reference points (with a down-sampling time interval  $\Delta T$ ) used for trajectory enhancement:  $T_i$  to  $T_{i+n-1}$  denote the  $n$  neighboring trajectory reference points for a generic timestamp  $T_0$ .

To fully utilize the relative (mainly from the INS unit) and absolute (mainly from the GNSS unit) trajectory information provided by the onboard navigation units, additional constraints related to the trajectory reference points are adopted. The first constraint is the minimization of correction differences between successive trajectory reference points, which ensures the smoothness and continuity of the refined trajectory, ultimately the quality of intra-dataset products. Furthermore, to ensure the absolute accuracy of the integrated imagery and LiDAR data, the trajectory reference point corrections can be forced to be close to zero, according to the expected accuracy of the initial trajectory. This ensures that the refined trajectory information does not significantly deviate from the initial values.

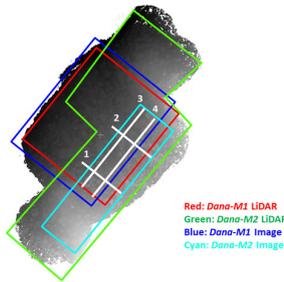
#### 4. EXPERIMENTAL RESULTS

In this section, the performance of the proposed UMSAT framework is qualitatively and quantitatively analyzed. Profiles are extracted from the LiDAR data as well as image-based object points from different systems and datasets for comparison. The accuracy of the camera geo-tagging information is further evaluated by back-projecting the features identified in LiDAR point cloud to imagery data. For quantitative evaluation, the root mean square (RMS) values of residuals for the camera/LiDAR observations to the respective object-space features are reported. Also, the normal distance between object-space feature pairs is reported if available.

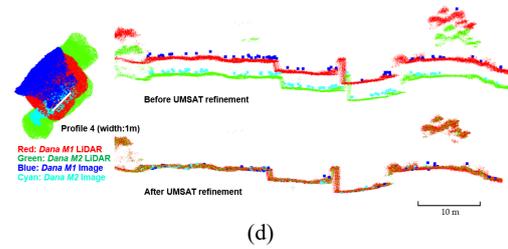
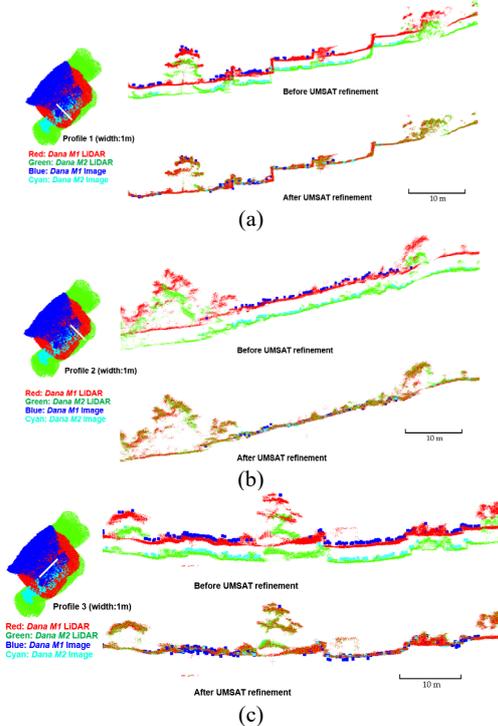
##### 4.1 Dana Island Datasets

For the in-house developed UAV system used for data acquisition at Dana Island, LiDAR system calibration parameters were derived from a rigorous system calibration and assumed to be errorless. In the integration process, LiDAR data from the first mission is used as control – i.e., trajectory from the *Dana-M1* dataset is fixed. In the LSA process, camera mounting parameters, trajectory information of the *Dana-M2* dataset, coordinates of image-based object points, as well as the parametric models of extracted planar patches are refined. Considering the dynamic characteristics of the UAV system, the trajectory of the second dataset is modeled by a 5 HZ reference points with a 2<sup>nd</sup>-order polynomial.

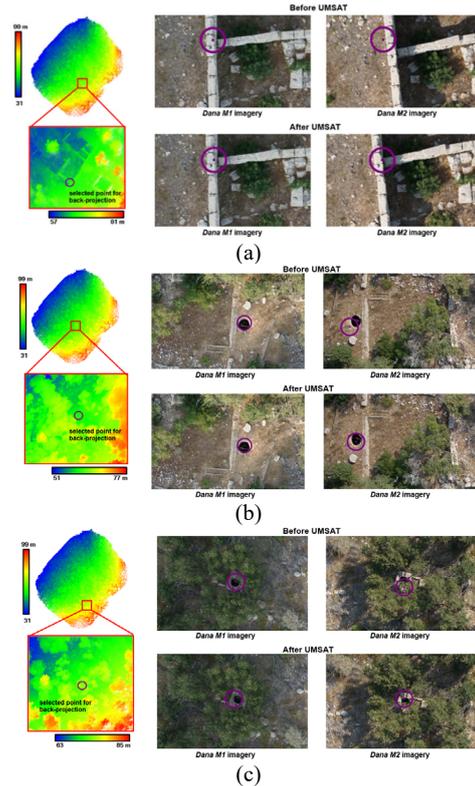
Four profiles from the overlapping areas of LiDAR data and image-based object points are extracted, as shown in Figure 9. It is worth mentioning that the UAV was operated under manual mode in the northern part of the *Dana-M2* dataset. Due to the limited overlap/side-lap among acquired images and the impact of dense vegetation, SfM could not derive image-based object points from that area. The alignment among camera and LiDAR data before and after the UMSAT optimization process are presented in Figure 10. By looking into the profiles before UMSAT, the intra-dataset alignment of camera and LiDAR is reasonable. However, large misalignment between the two datasets can be seen. By using the LiDAR data from the first mission as reference, camera and LiDAR data from both missions are well-aligned after the UMSAT throughout the overlapping areas, indicating that the integration process is successfully performed. The accuracy of the camera geo-tagging for the two missions is evaluated through backward projection, as shown in Figure 11. Three feature points are selected from the reference LiDAR data and back-projected onto the images (represented by magenta markers) from both datasets before and after UMSAT refinement. Similar findings can be observed from the back-projection results, indicating that initial intra-dataset camera/LiDAR alignment is reasonable.



**Figure 9.** Illustration of the LiDAR and image-based point cloud coverage from the two missions as well as the four profiles (represented by white lines) that will be used for evaluation superimposed on LiDAR data (colored by height).



**Figure 10.** Illustration of extracted profiles from the imagery and LiDAR point clouds from two datasets before/after the UMSAT refinement: (a) profile 1, (b) profile 2, (c) profile 3, and (d) profile 4.



**Figure 11.** Illustration of imagery back-projection accuracy before and after UMSAT refinement: selected feature point at (a) intersection of two walls and two cisterns in (b) and (c).

As for the quantitative evaluation of UMSAT results, RMS values of residuals for the camera/LiDAR constraints are evaluated, including (i) normal distance from LiDAR points to the respective object-space areal feature, (ii) back-projection error for image tie points, and (iii) normal distance from the image-based object point to corresponding areal feature. Table 3 lists the above metrics before and after the UMSAT optimization for the two datasets. It can be seen from this table that the initial misalignment in the range of 1-2 m is reduced to around 8 cm for the LiDAR areal features. An eight pixel back-projection error is achieved for the *Dana-M1* dataset while the error for the *Dana-M2* dataset is smaller (around 5 pixels). This is mainly because the trajectory for the latter is refined in the process. In terms of the RMS value of normal distances for image-based object point and LiDAR areal feature correspondence, these two types of features are in good agreement in the range of 2 cm after the UMSAT process.

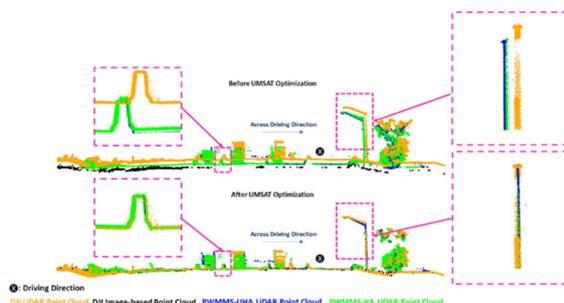
	RMS of normal distances for LiDAR areal features (cm)		RMS of back-projection errors (pixel)		RMS of normal distances of point-areal feature pair (cm)	
	before	after	before	after	before	after
Overall	145.181	8.293	53.847	8.069	141.404	2.125
Dana-M1	125.169	8.399	54.782	8.382	136.957	2.099
Dana-M2	159.763	8.204	45.943	5.003	159.243	2.234

**Table 3.** Quantitative evaluation of the before and after UMSAT optimization for the two Dana Island datasets.

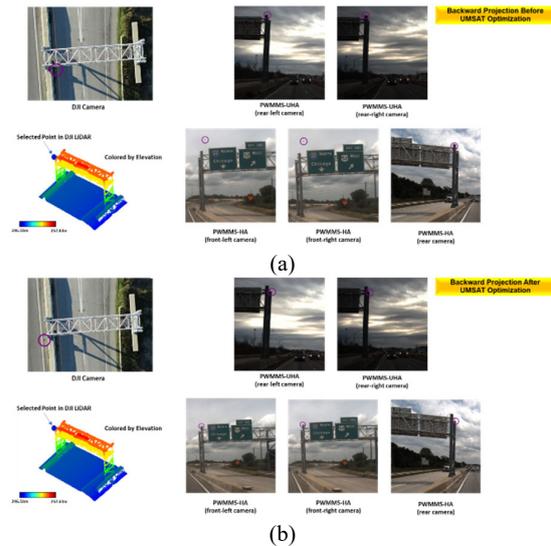
#### 4.2 I-65 Highway Datasets

For this study site, camera/LiDAR data collected by the UAV and wheel-based systems are integrated through the proposed UMSAT framework. LiDAR data acquired by DJI UAV is used as a reference due to the continuous access to GNSS signals during data acquisition. Specifically, the endpoints of linear features including lane markings and poles derived from the DJI UAV LiDAR are fixed as control information. For the wheeled MMS, camera/LiDAR IOP and mounting parameters have been established through a rigorous system calibration procedure. Therefore, their trajectory information will be refined through UMSAT using the extracted linear features. As for UAV imagery, only approximate geotagging information is available. While camera IOP were previously estimated in the SfM process, the EOP are estimated in UMSAT using both point and linear features. In this study, the trajectory of the wheeled systems is modeled by 1 HZ reference points with a 2<sup>nd</sup>-order polynomial. For the UAV images, we only solve for position/orientation corrections at their locations (i.e., a zero-order polynomial is used where a reference point is defined for each image).

To evaluate the performance of UMSAT, a profile perpendicular to the driving direction is extracted from the DJI UAV, PWMMS-UHA, and PWMMS-HA LiDAR data as well as the image-based point cloud from the DJI UAV camera, as shown in Figure 12. From the zoom-in window of the road and concrete traffic barrier, we can observe that the misalignment in the across driving and Z directions of multi-system LiDAR data is significantly reduced after the integration process. In addition, the close-up views of the light pole in the profile demonstrate the good alignment along the driving direction. As for the image-based point cloud, it is in agreement with the LiDAR data after the UMSAT refinement. After analyzing the agreement of camera and LiDAR in the object space, the camera geo-tagging accuracy is further assessed by back-projecting a point belonging to the top of a sign frame from the reference UAV LiDAR data onto imagery from all systems. As illustrated in Figure 13, reasonable back-projection accuracy is achieved for the datasets acquired from different dates.



**Figure 12.** Illustration of extracted profile from the LiDAR data of all systems as well as image-based point cloud from DJI UAV imagery before/after the UMSAT optimization.



**Figure 13.** Illustration of imagery back-projection accuracy (a) before and (b) after UMSAT refinement.

As for the quantitative evaluation of UMSAT results, RMS values of (i) normal distance from LiDAR points to the respective object-space linear feature, (ii)  $\alpha$  angle between the imaging ray for an intermediate point and the respective plane defined by the linear feature, and (iii) back-projection error for image tie points are evaluated. Table 4 lists these values before and after the UMSAT optimization for the DJI UAV, PWMMS-UHA, and PWMMS-HA datasets. For linear features, the initial misalignment is around 1.3 m for LiDAR data. Through the UMSAT process, this value decreases to 6-8 cm for the wheeled systems. In terms of image linear features, the RMS value of the  $\alpha$  angle is within 0.25° after the optimization. Moreover, a 1.3 pixel back-projection error is achieved for the DJI UAV camera data. In summary, qualitative and quantitative analysis suggests that the multi-temporal camera and LiDAR data are well-aligned through the proposed integration process.

	RMS of normal distances for LiDAR linear features (cm)		RMS of $\alpha$ angles for image linear features (°)		RMS of back-projection errors (pixel)	
	before	after	before	after	before	after
Overall	126.200	6.921	3.250	0.188	6.872	1.319
PWMMS-UHA	130.655	6.530	3.716	0.247	N/A	N/A
PWMMS-HA	122.169	7.246	5.249	0.220	N/A	N/A
DJI UAV Camera	N/A	N/A	0.915	0.140	6.872	1.319

**Table 4.** Quantitative evaluation of the before and after UMSAT optimization for the I-65 highway datasets.

#### 5. CONCLUSIONS

In this paper, a tightly-coupled camera/LiDAR integration workflow – unified multi-sensor advanced triangulation (UMSAT) – is proposed. This strategy can incorporate point/linear/areal features derived from imaging and ranging remote sensing systems, as well as trajectory information provided by GNSS/INS units. By minimizing discrepancies among conjugate features captured by different sensor modalities from either single or multiple systems, system calibration parameters and trajectory information are refined to ensure high positional quality of the derived camera and LiDAR

data/products. The proposed UMSAT framework is evaluated using camera/LiDAR data from two sample applications, namely archaeological mapping and transportation corridor documentation. The experimental results demonstrate that UMSAT can successfully align multi-temporal, multi-sensor, and multi-platform geospatial data in both cases, which is critical for future activities that require the fusion of camera/LiDAR data. The limitation of the proposed integration strategy is that GNSS/INS information is loosely-coupled within the process – i.e., GNSS/INS raw measurements are not included in the BA procedure. Therefore, future work will focus on developing a tightly-coupled GNSS/INS/camera/LiDAR integration process. Additionally, the study will explore automated feature extraction and matching procedures in different environments. Lastly, the feasibility of utilizing UMSAT in GNSS-challenging and potentially GNSS-denied environments will also be investigated.

#### ACKNOWLEDGEMENTS

The authors would like to acknowledge the technical/administrative/funding support from the Indiana Department of Transportation, Joint Transportation Research Program, the Civil Engineering Center for Applications of UAS for a Sustainable Environment (CE-CAUSE), and Information & Communication Technology- ICTWAY (Korea).

#### REFERENCES

- Ai, C., & Tsai, Y. J. (2016). An automated sign retroreflectivity condition evaluation methodology using mobile LIDAR and computer vision. *Transportation Research Part C: Emerging Technologies*, 63. <https://doi.org/10.1016/j.trc.2015.12.002>
- Caltagirone, L., Bellone, M., Svensson, L., & Wahde, M. (2019). LiDAR–camera fusion for road detection using fully convolutional neural networks. *Robotics and Autonomous Systems*, 111, 125–131. <https://doi.org/10.1016/j.robot.2018.11.002>
- Cheng, Y. T., Patel, A., Wen, C., Bullock, D., & Habib, A. (2020). Intensity thresholding and deep learning based lane marking extraction and lanewidth estimation from mobile light detection and ranging (LiDAR) point clouds. *Remote Sensing*, 12(9). <https://doi.org/10.3390/RS12091379>
- Furukawa, Y., & Ponce, J. (2009). Accurate camera calibration from multi-view stereo and bundle adjustment. *International Journal of Computer Vision*, 84(3), 257–268. <https://doi.org/10.1007/s11263-009-0232-2>
- Glira, P., Pfeifer, N., & Mandlbürger, G. (2016). Rigorous Strip adjustment of UAV-based laserscanning data including time-dependent correction of trajectory errors. *Photogrammetric Engineering & Remote Sensing*, 82(12), 945–954. <https://doi.org/10.14358/PERS.82.12.945>
- 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). <https://doi.org/10.3390/rs12030351>
- Kadhim, I., & Abed, F. M. (2021). The potential of lidar and uav-photogrammetric data analysis to interpret archaeological sites: A case study of chun castle in South-West England. *ISPRS International Journal of Geo-Information*, 10(1). <https://doi.org/10.3390/ijgi10010041>
- Li, J., Yang, B., Chen, C., & Habib, A. (2019). NRLI-UAV: Non-rigid registration of sequential raw laser scans and images for low-cost UAV LiDAR point cloud quality improvement. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158. <https://doi.org/10.1016/j.isprsjprs.2019.10.009>
- Moghadam, P., Bosse, M., & Zlot, R. (2013). Line-based extrinsic calibration of range and image sensors. *Proceedings - IEEE International Conference on Robotics and Automation*. <https://doi.org/10.1109/ICRA.2013.6631095>
- Schneider, S., Luettel, T., & Wuensche, H. J. (2013). Odometry-based online extrinsic sensor calibration. *IEEE International Conference on Intelligent Robots and Systems*. <https://doi.org/10.1109/IROS.2013.6696515>
- Shin, Y.-H., Shin, S.-Y., Rastiveis, H., Cheng, Y.-T., Zhou, T., Liu, J., Zhao, C., Varinlioglu, G., Rauh, N. K., Matei, S. A., & Habib, A. (2023). UAV-Based Remote Sensing for Detection and Visualization of Partially-Exposed Underground Structures in Complex Archaeological Sites. *Remote Sensing*, 15(7). <https://doi.org/10.3390/rs15071876>
- Verma, S., Berrio, J. S., Worrall, S., & Nebot, E. (2019). Automatic extrinsic calibration between a camera and a 3D Lidar using 3D point and plane correspondences. *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*. <https://doi.org/10.1109/ITSC.2019.8917108>
- Vilbig, J. M., Sagan, V., & Bodine, C. (2020). Archaeological surveying with airborne LiDAR and UAV photogrammetry: A comparative analysis at Cahokia Mounds. *Journal of Archaeological Science: Reports*, 33. <https://doi.org/10.1016/j.jasrep.2020.102509>
- Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179. <https://doi.org/10.1016/j.geomorph.2012.08.021>
- Zhou, L., & Deng, Z. (2014). LIDAR and vision-based real-time traffic sign detection and recognition algorithm for intelligent vehicle. *2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*. <https://doi.org/10.1109/ITSC.2014.6957752>
- Zhou, T., Hasheminasab, S. M., & Habib, A. (2021). Tightly-coupled camera/LiDAR integration for point cloud generation from GNSS/INS-assisted UAV mapping systems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 180. <https://doi.org/10.1016/j.isprsjprs.2021.08.020>