Unified Multi–Sensor Advanced Triangulation (UMSAT) for System Calibration and Trajectory Enhancement of Imaging and Ranging Sensors Onboard Mobile Mapping Systems

Tian ZHOU, Yi–Ting CHENG, Sang–Yeop SHIN, Young–Ha SHIN, Mona HODAEI, Jidong LIU, and Ayman HABIB, United States

Key words: mobile mapping systems, camera LiDAR integration, triangulation, system calibration, trajectory enhancement

SUMMARY

Mobile mapping systems (MMS) such as uncrewed aerial vehicles (UAVs) and wheeled platforms are widely used for a variety of applications such as precision agriculture, coastal monitoring, digital forestry, transportation management, infrastructure monitoring, bulk material estimation, and archaeological documentation. MMS are usually equipped with integrated global navigation satellite systems/inertial navigation systems (GNSS/INS) as well as imaging (e.g., RGB, multispectral, and hyperspectral cameras) and ranging (e.g., LiDAR) sensors. UAVs are becoming a viable alternative for small area mapping due to their ease of deployment, low cost, ability to fill an important gap between aerial and proximal mapping platforms, miniaturization/improvement of GNSS/INS georeferencing technologies, and proliferation of imaging/ranging sensors operating in different portions of the electromagnetic spectrum.

Integration of image and LiDAR data can provide a comprehensive 3D model of the area of interest. For such integration, ensuring a good alignment of derived products from single or several platforms is critical. Although many works have been conducted on this topic, there is still a need for a rigorous integration approach that minimizes the discrepancy between imagery and LiDAR data/products caused by inaccurate system calibration parameters and/or trajectory artifacts. This study proposes a tightly–coupled camera/LiDAR integration workflow for UAV and wheeled remote sensing systems aided by a GNSS/INS unit. More specifically, the paper presents a unified multi–sensor advanced triangulation (UMSAT), which can handle point, linear, and areal features derived from imaging and ranging remote sensing systems aided by GNSS/INS position and orientation unit. Through UMSAT, a general environment for system calibration and/or trajectory refinement will be explored for improving derived data/products from imaging and ranging remote sensing systems while focusing on transportation–related datasets. Experimental results from real datasets will be presented together with recommendations for future research to improve the performance of UMSAT in GNSS–challenging, and potentially GNSS–denied, environments.

Unified Multi–Sensor Advanced Triangulation (UMSAT) for System Calibration and Trajectory Enhancement of Imaging and Ranging Sensors Onboard Mobile Mapping Systems

Tian ZHOU, Yi–Ting CHENG, Sang–Yeop SHIN, Young–Ha SHIN, Mona HODAEI, Jidong LIU, and Ayman HABIB, United States

1. Introduction

Utilization of remote sensing technologies is becoming the norm for many applications due to their ability to map large areas in a short time at a reasonable cost. More specifically, the emergence of passive and active remote sensing modalities operating in different portions of the electromagnetic spectrum allows for the derivation of a rich set of information useful for various applications. Improvements and lower–cost of direct georeferencing technologies – i.e., integrated Global Navigation Satellite Systems/Inertial Navigation Systems (GNSS/INS) – enable control–free mapping. In spite of the improving capabilities of spaceborne and airborne remote sensing platforms, they do not provide reasonable spatial/temporal resolution at an affordable cost. Modern mobile mapping systems (MMS) – such as uncrewed aerial vehicles (UAVs) and wheeled systems – have emerged as promising platforms (Guan et al., 2014; Nex & Remondino, 2014). Using these systems is motivated by their low cost, ease of deployment, high maneuverability, and fine spatial/temporal resolution mapping. Recent availability of miniaturized sensing and direct georeferencing units facilitates the use of such platforms in a wide range of applications, such as precision agriculture, coastal monitoring, digital forestry, transportation management, infrastructure monitoring, and archaeological documentation.

Imaging sensors including RGB and multispectral/ hyperspectral (MS/HS) cameras are widely used. The majority of RGB cameras are based on a frame imaging mechanism, providing imagery in a two–dimensional (2D) raster data structure with spectral information. With recent developments in Structure from Motion (SfM) algorithms (Westoby et al., 2012), one can generate dense point clouds. However, this reconstruction is contingent on adequate overlap/side–lap among neighboring images and establishing sufficient matches. On the other hand, most HS cameras are based on push–broom technology – also known as line cameras – which capture 1D images at a time with fine frequency bands across the spectrum. Deriving 3D information from these cameras is difficult as there is no overlap between the captured 1D images (Hasheminasab et al., 2021). As for ranging sensors, Light Detection and Ranging (LiDAR) can directly provide 3D points with high geometric accuracy. Nevertheless, the lack of spectral/color information makes it difficult to derive semantic information for the acquired scene. Due to the complementary characteristics of imaging and ranging sensors, the integration of camera and LiDAR sensors can overcome their individual limitations, resulting in an accurate and better description of the object space. This integration will enhance the process of feature extraction, scene understanding, and visualization of derived products (Caltagirone et al., 2019).

Meaningful integration of multi–temporal data/products from different modalities onboard single or multiple systems is contingent on their positional quality. Accurate system calibration – including the sensor's interior orientation parameters (IOP) and mounting parameters relating the sensors to the INS' Inertial Measurement Unit (IMU) body frame – and trajectory information are essential for ensuring the positional accuracy. 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 extracted from both modalities through a Least Squares Adjustment (LSA) procedure. Depending on the type of utilized features, calibration techniques can be categorized into target–based and target–less approaches. For example, Zhang and Pless (2004) used a planar checkerboard for establishing the mounting parameters relating camera and 2D LiDAR units. Several studies extended this work for calibrating systems equipped with camera and 3D LiDAR sensors (Mirzaei et al., 2012; Ravi et al., 2018; Verma et al., 2019). In addition to point–to–plane geometric constraints, other feature correspondences such as line–to–plane (Zhou, 2014) and point–to–point (Beltran et al., 2022) have been also adopted using custom–built targets. However, these approaches are time–consuming and might not be practical when frequent system calibration is required. Early works dealing with in–situ calibration using target–less strategies were based on manual identification of conjugate natural points and linear features in indoor scenes (Moghadam et al., 2013; Scaramuzza et al., 2007). 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 (Schneider et al., 2013) or SfM (Glira et al., 2016; Li et al., 2019; Zhou et al., 2021) to establish conjugate features in image and LiDAR data. More specifically, these techniques rely on deriving image–based point clouds and then matching those 3D points to LiDAR–derived features. Trajectory information is usually refined in these approaches to achieve the best alignment between camera and LiDAR data.

Despite that extensive body of work that has been conducted, there is still a need for a rigorous integration approach that minimizes the discrepancy between camera and LiDAR data/products caused by inaccurate system calibration parameters and/or trajectory artifacts. This study proposes a semi–automated, tightly–coupled camera/LiDAR integration workflow for UAV and wheeled remote sensing systems aided by a GNSS/INS unit. More specifically, the paper presents a unified multi–sensor advanced triangulation (UMSAT), which can handle point, linear, and areal features derived from imaging (e.g., frame cameras and push–broom scanners) and ranging modalities aided by GNSS/INS position and orientation unit.

2. Methodology

The success of any multi–modal geospatial data integration activity is contingent on ensuring the positional quality of such data (e.g., proper georeferencing of the used sensors together with comprehensive modeling of the point positioning equations relating their measurements to the respective ground coordinates). Before introducing the proposed UMSAT framework, the point positioning models are first introduced. In general, establishing the point positioning equations for either LiDAR or imaging systems proceeds in two steps. First, we need to define the laser beam or imaging ray relative to the sensor coordinate system. This definition is based on the sensor measurements (i.e., laser range/pointing direction for a LiDAR and image coordinate measurements for a camera) together with the IOP of the used sensor (i.e., parameters describing the encoder mechanism for a LiDAR or principal point coordinates, principal distance, and distortion parameters for a camera). Second, the position and orientation of the laser beam or imaging ray relative to the mapping frame are established through the Exterior Orientation Parameters (EOP) that describe the position and orientation of the sensor relative to the mapping frame. For a GNSS/INS–assisted system, the EOP are derived using the post–processed GNSS/INS trajectory and mounting parameters relating these sensors to the corresponding IMU body frame.

The point positioning models for LiDAR and frame/line camera units are described in Equations (1)–(3), respectively. These models are also graphically explained in Figure 1. In Equation (1), $r_I^{lu(t)}$ denotes the position of the footprint of a laser beam, emitted at time t, relative to the laser unit frame; $r_{b(t)}^m$ and $R_{b(t)}^m$ are the position and orientation of the IMU body frame relative to the mapping frame at time t; r_{lu}^b and R_{lu}^b represent the lever arm and boresight rotation matrix relating the laser unit and IMU body frame coordinate systems. The derivation of $r_I^{lu(t)}$ is based on the range/pointing direction measurements of the LiDAR unit as well as its IOP. For the point positioning for frame and line imaging systems (Equations (2) and (3)), $r_i^{cf(t)}$ and $r_i^{cl(t)}$ represent the imaging rays for point i relative to the frame/line camera coordinate systems at time t. This term is derived from the image coordinates of point i (x_i and y_i) and camera IOP, including the principal point coordinates of 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 $dist_{y_i}$). The main difference between frame and line cameras is that while both x, y components of image coordinates for the former have variable values depending on image point location, for the latter, the y coordinates are always constant – e.g., y_i = 0 for systems with the scan line vertically below the camera perspective center. r_{cf}^b/R_{cf}^b and r_{cl}^b/R_{cl}^b represent the lever arm and boresight rotation matrix relating the frame camera/line camera and IMU body frame coordinate systems. Different from LiDAR, image–based 3D reconstruction involves an unknown scale factor ($\lambda(i,cf,t)/\lambda(i,cl,t)$ for image point i captured by frame camera cf or line camera cl at time t), which needs to be estimated.

$$r_I^m = r_{b(t)}^m + R_{b(t)}^m r_{lu}^b + R_{b(t)}^m R_{lu}^b r_I^{lu(t)} \tag{1}$$

$$r_I^m = r_{b(t)}^m + R_{b(t)}^m r_{cf}^b + \lambda(i,cf,t) R_{b(t)}^m R_{cf}^b r_i^{cf(t)}, r_i^{cf(t)} = \begin{bmatrix} x_i - x_p - dist_{x_i} \\ y_i - y_p - dist_{y_i} \\ -f \end{bmatrix} \tag{2}$$

$$r_I^m = r_{b(t)}^m + R_{b(t)}^m r_{cl}^b + \lambda(i,cl,t) R_{b(t)}^m R_{cl}^b r_i^{cl(t)}, r_i^{cl(t)} = \begin{bmatrix} x_i - x_p - dist_{x_i} \\ 0 - y_p - dist_{y_i} \\ -f \end{bmatrix} \tag{3}$$

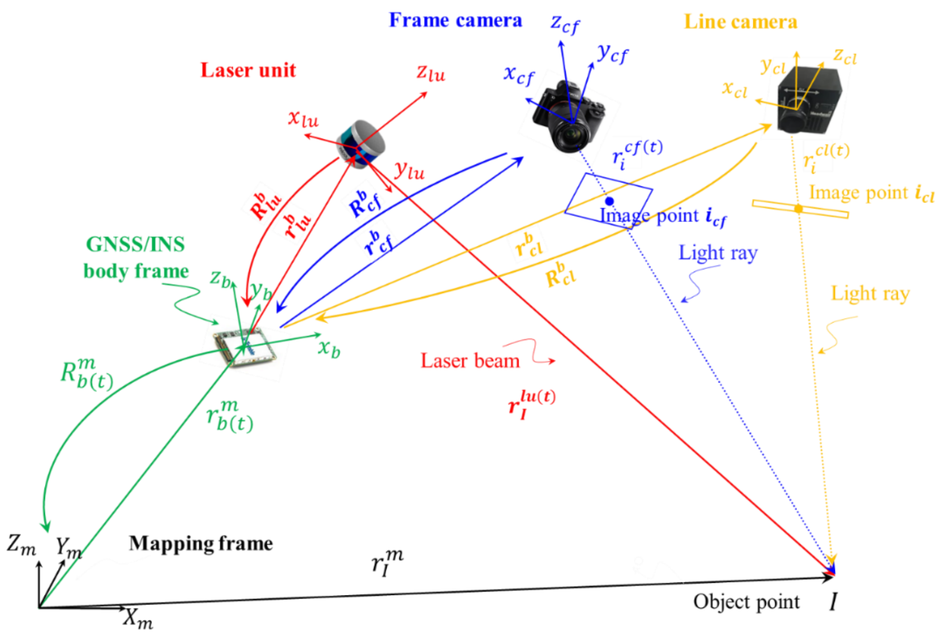


Figure 1. Schematic diagram of the point positioning principle for LiDAR and frame/line cameras units onboard a GNSS/INS–assisted MMS.

From the LiDAR/image–based point positioning equations, it is evident that accurate trajectory information and system calibration parameters (including sensor IOP and mounting parameters) are critical for producing properly georeferenced data from LiDAR and imaging systems. Therefore, to improve the positional quality of acquired data, a system–driven triangulation strategy is proposed. As illustrated in Figure 2, the triangulation aims at minimizing discrepancies among conjugate features (including point, linear, and areal features) captured by different sensor modalities from single/multiple systems through what will be denoted as universal multi–sensor advanced triangulation (UMSAT).

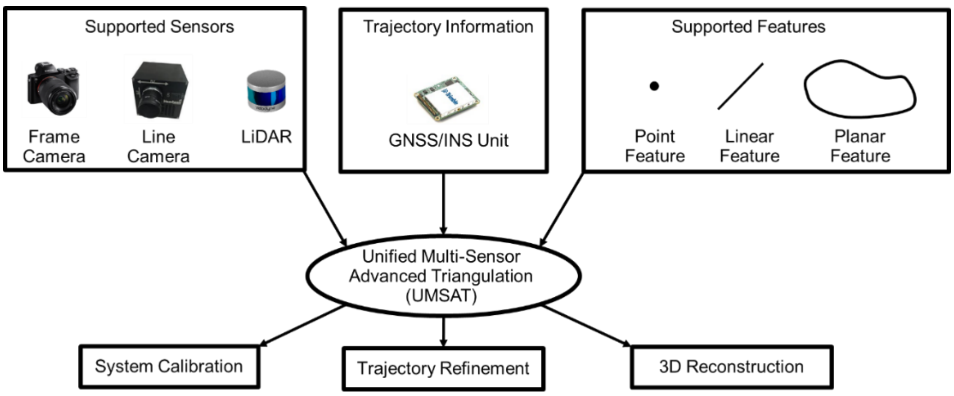


Figure 2. Schematic diagram of the functionality of the proposed UMSAT.

Point features are mainly adopted for cameras. The derivation of image–based object points can be conducted through SfM algorithms. For an object point *I* and its corresponding conjugate image points, the back–projection error is adopted as a cost function in UMSAT. Using the frame camera as an example, the point positioning equation can be reformulated into Equation (4); to eliminate the unknown scale factor from this equation, the first and second rows are divided by the third one, and the image point coordinates are expressed as Equation (5). Based on this equation, the differences between observed image coordinates and predicted ones using estimated unknowns (i.e., back–projection errors) are minimized in the LSA process.

$$r_i^{cf(t)} = \frac{1}{\lambda(i,cf,t)}\left[R_b^{cf} R_m^{b(t)}\left(r_I^m - r_{b(t)}^m - R_{b(t)}^m r_{cf}^b\right)\right] = \frac{1}{\lambda(i,cf,t)}\begin{bmatrix} N_x \\ N_y \\ D \end{bmatrix} \tag{4}$$

$$\begin{aligned} x_i &= -c\frac{N_x}{D} + x_p + dist_{x_i} \\ y_i &= -c\frac{N_y}{D} + y_p + dist_{y_i} \end{aligned} \tag{5}$$

In terms of linear features, they can be extracted from images and LiDAR data. For imagery, the Line Segment Detector (LSD) (Grompone Von Gioi et al., 2010) can be used to automatically derive image linear features. For LiDAR data, linear features could be derived through a segmentation process followed by Principal Component Analysis (PCA) analysis. In UMSAT, an object space linear feature is represented by two endpoints P_A and P_B. Two optimization target functions are implemented for minimizing the discrepancies between conjugate image/LiDAR lines. The involved quantities in the two target functions are schematically illustrated in Figure 3. The LiDAR optimization target function minimizes the normal distance between the mapping coordinates of a LiDAR point I and the linear feature it belongs to (defined by the two endpoints P_A and P_B). This constraint is mathematically described in Equation (6), where $\|x\|$ represents the L_2–norm of the vector x, $r_{P_A}^m$ and $r_{P_B}^m$ are the object coordinates of the two endpoints, and r_I^m is the coordinates of LiDAR point I. The image optimization target function forces the vector from the perspective center (PC) to an 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 $r_{PC-P_A}^m$ and $r_{PC-P_B}^m$, as shown in Figure 3). This constraint is mathematically presented by the triple product in Equation (7), where $r_{PC-P_A}^m$ is defined by $r_{P_A}^m - r_{c(t)}^m$ with $r_{c(t)}^m$ representing the camera position relative to the mapping frame at time t; and r_{PC-i}^m is the vector from the camera perspective center to an intermediate image point i along the line in the mapping frame as represented as $R_{b(t)}^m R_c^b r_i^{c(t)}$. To analyze the residual of such constraint following the LSA, the angle α between the vector r_{PC-i}^m and the plane defined by the PC and object line endpoints P_A/P_B (as shown in Figure 3) is evaluated.

Areal features are only used for LiDAR sensors as they cannot provide redundant information for imagery. Areal features can be automatically extracted from LiDAR data through various approaches. In UMSAT, areal features are modeled as planes. The respective target function minimizes the normal distance between the LiDAR point I to the areal feature it belongs to, as mathematically described in Equation (8). In this equation, A, B, C, D are the plane parameters and (X_I, Y_I, Z_I) are the coordinates of LiDAR point I in the mapping frame.

$$\frac{\left\|(r_{P_B}^m - r_{P_A}^m) \times (r_{P_B}^m - r_I^m)\right\|}{\left\|r_{P_B}^m - r_{P_A}^m\right\|} = 0 \tag{6}$$

$$(r_{P.C.-A}^m \times r_{P.C.-B}^m) \cdot r_{P.C.-i}^m = \left(\left[r_{P_A}^m - \left[r_{b(t)}^m + R_{b(t)}^m r_c^b\right]\right] \times \left[r_{P_B}^m - \left[r_{b(t)}^m + R_{b(t)}^m r_c^b\right]\right]\right) \cdot R_{b(t)}^m R_c^b r_i^{c(t)} = 0 \tag{7}$$

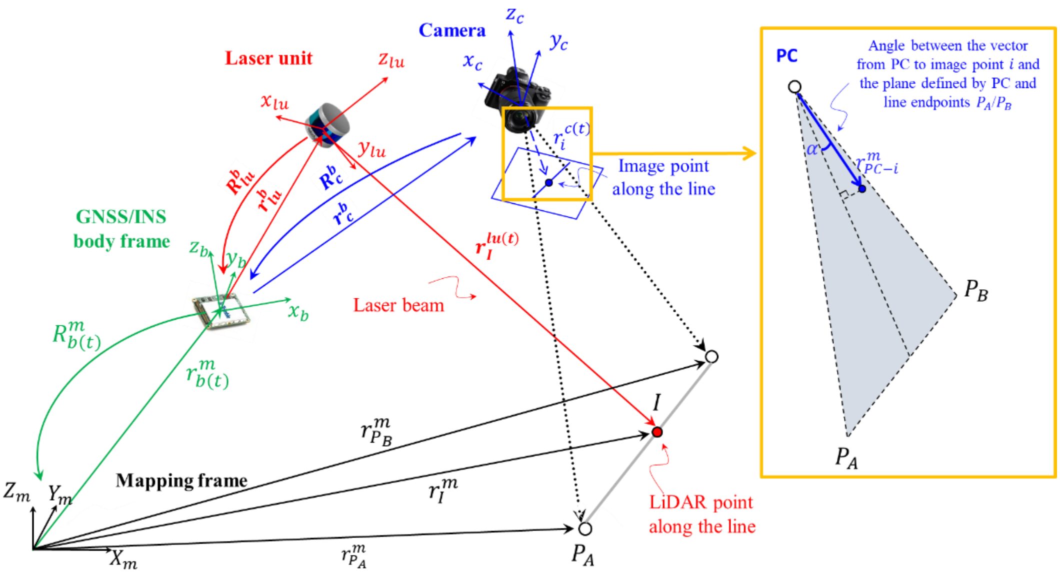


Figure 3. Schematic illustration of image/LiDAR points along the linear feature for the SMART system (points I and i represent the points along the linear feature observed by LiDAR and camera, respectively).

$$AX_I + BY_I + CZ_I + D = 0 \tag{8}$$

The involved parameters in the above optimization functions include the respective sensor's IOP and mounting parameters, trajectory information at the timestamp of the observation, and parameters for the respective object point, linear, or areal feature. In this study, trajectory information is refined by estimating corrections $\left(\delta r_{b(t)}^m, \delta R_{b(t)}^m\right)$ to the original position/orientation parameters derived from the post–processed GNSS/INS observations. Solving for the trajectory corrections at the timestamp for every camera/LiDAR observation is not recommended as it would cause over–parametrization in the LSA. Since we are dealing with a platform that has a relatively smooth trajectory with moderate dynamics, the original high–frequency (e.g., 100–200 Hz) trajectory is down–sampled (i.e., using a down–sampling time interval ΔT). The down–sampled trajectory points are henceforth denoted as trajectory reference points, as shown in Figure 4. 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. Symbolically, this polynomial modeling is expressed in Equation (9), where it can be seen that for a generic timestamp, T_0, its trajectory corrections (denoted generically 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 chosen based on the nature of platform dynamics.

$$\delta\theta_{b(T_0)}^{m} = f\left(T_0, T_i, T_{i+1}, \dots, T_{i+n-1}, \delta\theta_{b(T_i)}^{m}, \delta\theta_{b(T_{i+1})}^{m}, \dots, \delta\theta_{b(T_{i+n-1})}^{m} \right) \quad (9)$$

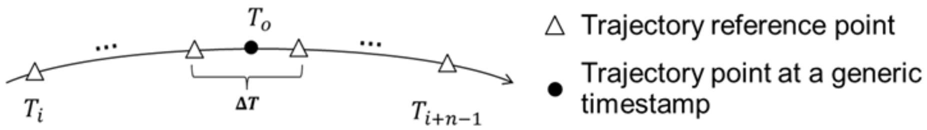


Figure 4. Down–sampled trajectory reference points (with down–sampling time interval ΔT) used for trajectory enhancement: T_i to T_{i+n-1} denote the n neighboring trajectory reference points for a generic timestamp T_0.

Other than the above target functions, other constraints are also included in the UMSAT. To guarantee the smoothness/continuity of the refined trajectory, correction differences for successive trajectory reference points are minimized. For an image–based object point P that belongs to an areal LiDAR feature, the normal distance between P and the corresponding plane can be also minimized in the LSA. Similarly, for a linear–areal feature correspondence, the endpoints to plane distances can be minimized. These correspondences are useful to integrate various camera and LiDAR features. Finally, UMSAT supports the incorporation of prior information for the unknowns including system calibration parameters, trajectory information, and/or feature parameters. This is conducted through the incorporation of pseudo observations.

3. Utilized Mobile Mapping Systems and Datasets Description

3.1. Wheeled and UAV Mobile Mapping Systems

This study involves 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). The PWMMS–UHA, as displayed in Figure 5a, is equipped with two single–beam LiDAR scanners: Riegl VUX 1HA and Z+F Profiler 9012. These scanners deliver a $360°$ horizontal field of view (FOV). Each scanner can deliver upto 1,000,000 points per second. Two rear–facing FLIR Flea2 FireWire cameras are installed on the PWMMS–UHA. Both cameras have a maximum image resolution of 5.0 MP and are synchronized to capture images at a rate of 1 frame every 0.75 s. All sensors are directly georeferenced by a NovAtel ProPak6 and ISA–100C GNSS/INS unit. The PWMMS–HA, as shown in Figure 5b, includes four multi–beam LiDAR scanners: three Velodyne HDL–32Es and one Velodyne VLP–16 Hi–Res. The HDL–32E consists of 32 radially oriented laser rangefinders aligned vertically from $-30.67°$ to $+10.67°$. The VLP–16 Hi–Res has 16 radially oriented laser rangefinders with a $20°$ vertical FOV. The point capture rates for HDL–32E and VLP–16 Hi–Res are 700,000 and 300,000 points per second, respectively. Three FLIR Grasshopper3 9.1MP GigE cameras are also mounted on the PWMMS–HA: two forward–facing and one rear–facing. The cameras are synchronized to capture one 1 frame per second per camera. The PWMMS–HA sensors are directly georeferenced by an Applanix POS LV 220

GNSS/INS unit. In addition to wheeled MMS, an off–the–shelf UAV system, a DJI M300 equipped with the Zenmuse L1 LiDAR sensor is used (Figure 5c). The Zenmuse L1 integrates a Livox LiDAR module, a camera, and an IMU on a 3–axis stabilized gimbal. The LiDAR horizontal and vertical FOVs are 70.4° and 4.5°, respectively. The point capture rate is 700,000 points per second. 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, a position accuracy of ±1 to ±1.5 cm and attitude accuracy of ±0.025° and ±0.15° for pitch/roll and heading, respectively, can be achieved.

Figure 5. Illustrations of (a) Purdue wheel–based MMS—ultra high accuracy system (PWMMS–UHA), (b) Purdue wheel–based MMS—high accuracy system (PWMMS–HA), and (c) DJI M300 UAV equipped with the Zenmuse L1 (adapted from DJI website).

3.2 Study Site and Datasets Description

In this study, the image/LiDAR data were acquired along a 0.5 mile segment of the I–65 highway close to Lebanon, IN, United States, as shown in Figure 6. The site is rich with point, linear, and areal features, which could be identified in the imagery and LiDAR data. A total of three datasets are used.

Table 1 provides a summary of the relevant information for the datasets acquired from the PWMMS–UHA, PWMMS–HA, and DJI systems.

Figure 6. Location of study site adopted in this research.

Table 1. Specifications of acquired datasets for the wheeled and UAV systems.

Platform	Data Acquisition Date	Driving/flight Speed	Number of Collected LiDAR Points (in millions)	Number of Captured Images
PWMMS–UHA	2022.07.10	~50 mph	~23	52
PWMMS–HA	2022.12.02	~50 mph	~63	158
DJI UAV*	2022.08.02	~8.5 mph	~85	88

* Above ground flying height is 70 meters.

4. Experimental Results

In this study, camera/LiDAR data collected by the three systems are integrated through the proposed UMSAT framework. LiDAR data acquired by the DJI system is used as a reference due to the high positional accuracy provided through uninterrupted GNSS–signal. For the captured UAV images, we have only access to approximate geotagging information. While camera IOP were derived through an SfM process, the EOP are estimated through UMSAT. For the wheeled MMS, rigorous system calibration has been conducted. Although onboard GNSS/INS units provide trajectory with reasonable accuracy, misalignments can be observed due to GNSS–signal interruptions for the wheeled MMS units. Therefore, system calibration parameters (IOP and mounting parameters for the camera and LiDAR units) of PWMMS–UHA and PWMMS–HA are fixed in the LSA while refining their trajectory through UMSAT.

For the triangulation features, lane markings are modeled as linear features, as shown in Figure 7. Specifically, skip–lines are modeled as individual linear features. Edge lines, on the other hand, are divided into short straight line segments. A geometry–based approach is adopted to extract lane markings from the UAV and wheeled LiDAR data (Cheng et al., 2020). Then, lane markings derived from PWMMS–UHA, PWMMS–HA, and DJI LiDAR are matched according to their spatial proximity. Corresponding image lines are established through back–projection of LiDAR linear features onto the imagery. Since lane markings are almost parallel, they only provide control in the vertical and across driving directions. Therefore, four poles (colored in red in Figure 7) are manually extracted from the LiDAR data to provide control along the driving direction. These poles are modeled as linear features as well. For the UAV camera imagery, other than the manually established linear features, a SfM is conducted to derive image tie points with the corresponding object points depicted in Figure 7. Moreover, intermediate points along the lane markings and poles have been also established in the imagery.

In this study, the endpoints of line segments (including lane markings and poles) from the DJI UAV LiDAR are used as control information. The PWMMS–UHA (with two LiDAR units and two cameras), PWMMS–HA (with four LiDAR units and three cameras), and UAV camera datasets are included by considering both point and linear features. For the two wheeled systems, the trajectory is modeled by 1HZ reference points with a 2nd order polynomial. For the UAV images, we only solve for position/orientation corrections at their locations (i.e., a zero–order polynomial is used – i.e., a reference point is defined for each image).

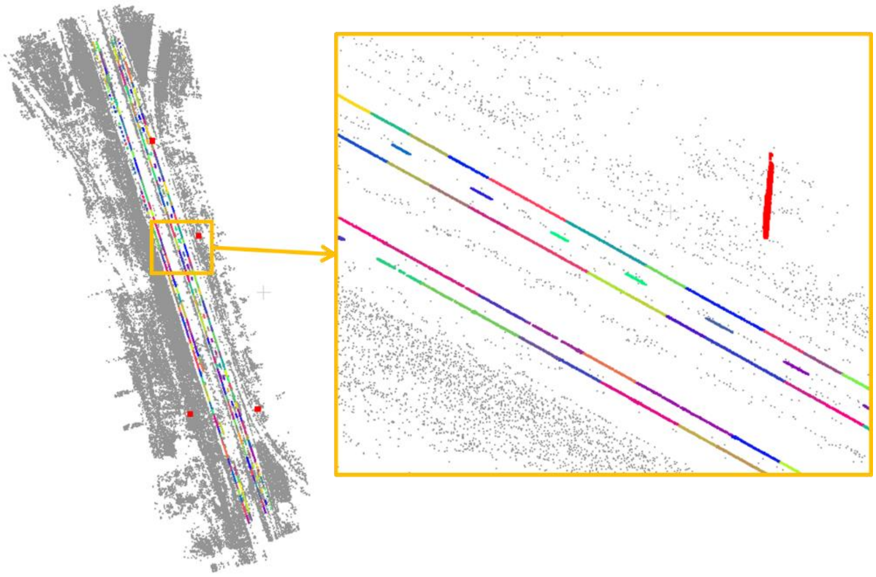


Figure 7. Established lane markings (randomly colored by the feature ID) and four poles (colored in red) from the DJI LiDAR data, as well as the image–based object points (colored in grey) derived from SfM processing of the DJI imagery.

The performance of UMSAT is first evaluated through qualitative analysis. Specifically, profiles perpendicular to the driving direction are extracted from the DJI, PWMMS–UHA, and PWMMS–HA LiDAR data as well as the image–based point cloud from the DJI camera (Figure 8). We can observe that the misalignment in the across driving and Z directions is minimized after the UMSAT process. To evaluate the alignment along the driving direction, one of the light poles is extracted from the profiles and also shown in Figure 8, where we can see good alignment along the driving direction (refer to the close–up views in Figure 8).

Having examined the 3D alignment between the LiDAR and image–based point clouds, the accuracy of the camera geo–tagging for all systems is evaluated through backward projection, as shown in Figure 9. A point on one of the light poles was selected in the DJI LiDAR data and back–projected onto the images from DJI, PWMMS–HA, and PWMMS–UHA systems. As shown in Figure 9b, the 2D misalignment is minimized following the UMAST optimization.

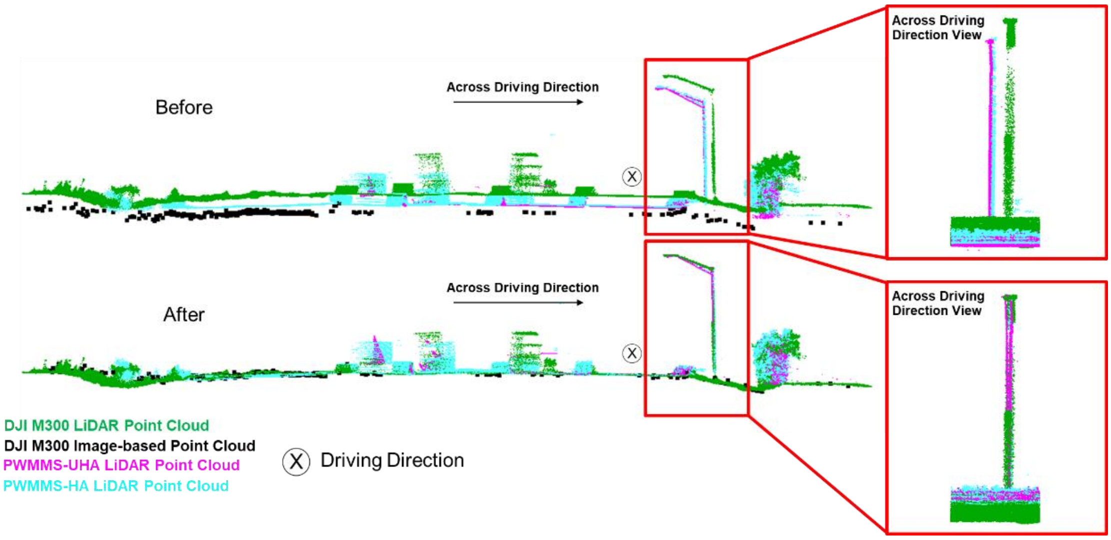
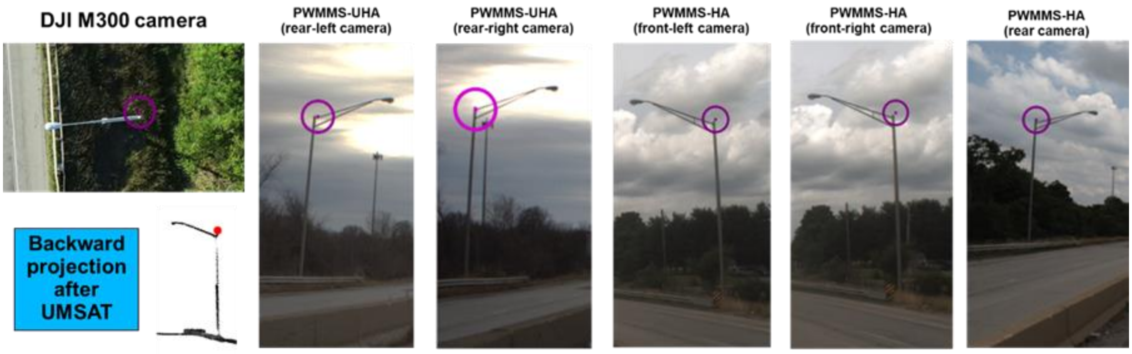


Figure 8. Illustration of extracted profiles/features from the LiDAR data as well as image–based point cloud before/after UMSAT optimization showing improved alignment in the across/along driving and vertical directions.

(a)



(b)

Figure 9. Illustration of image back–projection accuracy (a) before and (b) after applying UMSAT.

The quantitative evaluation of UMSAT is analyzed through the Root Mean Square (RMS) values of residuals for the camera/LiDAR constraints, including (i) normal distance from LiDAR points to the respective object–space linear features, (ii) α angle between the imaging ray for an intermediate point and the respective plane through the linear feature, and (iii) back–projection error for image tie points. Table 2 lists the above metrics before and after the UMSAT optimization for the different systems. It can be seen from this table that the before optimization misalignment, in the range of 1.2 – 1.3 m for the LiDAR linear features, is reduced to roughly 7 cm. As for the image linear features, the RMS of the α angle is reduced to almost 0.2°. As for the image tie points, a 1.3 pixel back–projection error is achieved for the DJI camera data.

Table 2. Quantitative evaluation of the pre/post–UMSAT optimization for the three datasets.

	RMS of normal distances for the LiDAR linear features (m)		RMS of α angles for the image linear features (degree)		RMS of back–projection erros for the image tie points (pixel)	
	Before	After	Before	After	Before	After
Overall	1.262	0.069	3.250	0.191	6.872	1.319
PWMMS–UHA	1.307	0.065	3.716	0.269	N/A	N/A
PWMMS–HA	1.222	0.073	5.250	0.213	N/A	N/A
DJI Camera	N/A	N/A	0.915	0.140	6.872	1.319

5. Conclusions and Recommendations for Future Work

This study presented a unified multi–sensor advanced triangulation (UMSAT), which can handle point, linear, and areal features derived from imaging and ranging remote sensing systems aided by a GNSS/INS position and orientation unit. Camera/LiDAR data collected by two wheeled and one UAV MMS over a highway study site are used to evaluate the performance of the proposed strategy. Experimental results indicate that through UMSAT, camera and LiDAR data from these systems are well–aligned (i.e., indicating that multi–temporal, multi–sensor, and multi–platform geospatial data are ready for subsequent integration activities). The current limitation of the proposed strategy is the time–consuming, manual measurements of image linear features. Therefore, automated feature extraction and matching procedures from camera and LiDAR data will be explored. Moreover, the feasibility of using UMSAT in GNSS–challenging, and potentially GNSS–denied, environments will be investigated.

REFERENCES

Beltran, J., Guindel, C., De La Escalera, A., & Garcia, F. (2022). Automatic Extrinsic Calibration Method for LiDAR and Camera Sensor Setups. *IEEE Transactions on Intelligent Transportation Systems*, *23*(10). https://doi.org/10.1109/TITS.2022.3155228

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

Glira, P., Pfeifer, N., & Mandlburger, 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

Grompone Von Gioi, R., Jakubowicz, J., Morel, J. M., & Randall, G. (2010). LSD: A fast line segment detector with a false detection control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(4). https://doi.org/10.1109/TPAMI.2008.300

Guan, H., Li, J., Yu, Y., Wang, C., Chapman, M., & Yang, B. (2014). Using mobile laser scanning data for automated extraction of road markings. *ISPRS Journal of Photogrammetry and Remote Sensing*, *87*. https://doi.org/10.1016/j.isprsjprs.2013.11.005

Hasheminasab, S. M., Zhou, T., Laforest, L., & Habib, A. (2021). Multiscale Image Matching for Automated Calibration of UAV–Based Frame and Line Camera Systems. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*. https://doi.org/10.1109/JSTARS.2021.3062573

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

Mirzaei, F. M., Kottas, D. G., & Roumeliotis, S. I. (2012). 3D LiDAR–camera intrinsic and extrinsic calibration: Identifiability and analytical least–squares–based initialization. *The International Journal of Robotics Research*, *31*(4), 452–467. https://doi.org/10.1177/0278364911435689

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

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: A review. In *Applied Geomatics* (Vol. 6, Issue 1). https://doi.org/10.1007/s12518–013–0120–x

Ravi, R., Lin, Y.–J., Elbahnasawy, M., Shamseldin, T., & Habib, A. (2018). Simultaneous system calibration of a multi–LiDAR multicamera mobile mapping platform. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(5), 1694–1714. https://doi.org/10.1109/JSTARS.2018.2812796

Scaramuzza, D., Harati, A., & Siegwart, R. (2007). Extrinsic self calibration of a camera and a 3D laser range finder from natural scenes. *IEEE International Conference on Intelligent Robots and Systems*. https://doi.org/10.1109/IROS.2007.4399276

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

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

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

Zhang, Q., & Pless, R. (2004). Extrinsic calibration of a camera and laser range finder (improves camera calibration). *IROS*, *3*, 2301–2306. https://doi.org/10.1109/IROS.2004.1389752

Zhou, L. (2014). A new minimal solution for the extrinsic calibration of a 2D LIDAR and a camera using three plane–line correspondences. *IEEE Sensors Journal*, *14*(2). https://doi.org/10.1109/JSEN.2013.2284789

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

BIOGRAPHICAL NOTES

Tian Zhou, Yi–Ting Cheng, Sang–Yeop Shin, Young–Ha Shin, Mona Hodaei, and Jidong Liu are graduate student members of the Digital Photogrammetry Research Group (DPRG) at the Lyles School of Civil Engineering, Purdue University. The DPRG group is led by Ayman Habib.

CONTACTS

Ayman Habib
Organization: Purdue University
Address: 550 Stadium Mall Drive, West Lafayette, IN 47907
City: West Lafayette
COUNTRY: United States
Tel. +1 (765) 496–0173
Email: ahabib@purdue.edu
Web site: https://engineering.purdue.edu/CE/Academics/Groups/Geomatics/DPRG

ACKNOWLEDGEMENT

The authors would like to acknowledge the technical/administrative/funding support from the Indiana Department of Transportation (INDOT – USA), Joint Transportation Research Program (JTRP – USA), and Information & Communication Technology – ICTWay (Korea).