Combining Nadir, Oblique, and Façade Imagery Enhances Reconstruction of Rock Formations Using Unmanned Aerial Vehicles

Yu-Hsuan Tu, *Member, IEEE*, Kasper Johansen, Bruno Aragon, Bonny M. Stutsel,
Yoseline Ángel, Omar A. López Camargo, Samir K. M. Al-Mashharawi, Jiale Jiang,
Matteo G. Ziliani, and Matthew F. McCabe

Abstract—Developments in computer vision, such as structure from motion and multiview stereo reconstruction, have enabled a range of photogrammetric applications using unmanned aerial vehicles (UAV)-based imagery. However, some specific cases still present reconstruction challenges, including survey areas composed of steep, overhanging, or vertical rock formations. Here, the suitability and geometric accuracy of four UAV-based image acquisition and data processing scenarios for topographic surveying applications in complex terrain are assessed and compared. The specific cases include the use of: 1) nadir imagery; 2) nadir and oblique imagery; 3) nadir and façade imagery; and 4) nadir, oblique, and façade imagery to reconstruct a topographically complex natural surface. Results illustrate that including oblique and façade imagery to supplement the more traditional nadir collections significantly improves the geometric accuracy of point cloud data reconstruction by approximately 35% when assessed against terrestrial laser scanning data of near-vertical rock walls. Most points (99.41%) had distance errors of less than 50 cm between the point clouds derived from the nadir imagery and nadir–oblique–façade imagery. Apart from delivering enhanced spatial resolution in façade details, the geometric accuracy improvements achieved from integrating nadir, oblique, and façade imagery provide value for a range of applications, including geotechnical and geohazard investigations. Such gains are particularly relevant for studies assessing rock integrity and stability, and engineering design, planning, and construction, where information on the position of rock cracks, joints, faults, shears, and bedding planes may be required.

Index Terms—Computer vision, photogrammetry, structure from motion (SfM), topographic survey, unmanned aerial vehicles (UAV).

I. INTRODUCTION

TOPOGRAPHIC surveys are essential elements in supporting a range of civil engineering and geomorphological applications, including studies of erosion deposition and transport, structural stability and movement, and geohazard investigations [1]–[3]. Two decades ago, topographic

data were predominantly collected using ground-based point sampling methods from differential global navigation satellite systems (GNSS) and total stations [4]. Since the introduction of laser scanning technologies, light detection and ranging (LiDAR) systems have become widely used to acquire topographic data [5]. However, both terrestrial and airborne LiDAR systems are expensive for repeated survey tasks, which are often required for applications where temporal dynamics are of interest. Furthermore, while terrestrial laser scanning (TLS) can provide very high spatial resolution surface models of high accuracy, the spatial coverage can be limited and also restricted by line-of-sight viewing requirements [6], [7]. Area accessibility can also pose a key constraint, particularly in hazardous or complex terrains. Traditional airborne laser scanning (ALS) can cover large areas within a relatively short time, with data acquisition specifications suited for detailed terrain analysis. On the other hand, ALS systems need significant expertise to overcome the complexity of the operation and still require line-of-sight viewing, and the cost of operation and equipment is usually much higher than for TLS operations for local-scale surveying tasks [6], [8].

Digital photogrammetry using optical remote sensing presents as a relatively cost-effective alternative to acquire topographic data. Using well-established techniques of multiview stereopsis (MVS), the elevation of the Earth's surface can be estimated [9]. Factors such as image spatial resolution, sampling density, interpolation methods, and optical sensor models can significantly affect the accuracy of the produced topographic data [9]–[11]. Therefore, high-quality image capture, rigorous analysis, and user-expertise are all required in the production of accurate topographic data, which have limited data provision to specialized consulting services in the past [8], [12]. In recent years, there has been dramatic growth in the use of low-altitude remote sensing platforms, particularly through the application of unmanned aerial vehicles (UAV) [13]. The installation of commercial-off-the-shelf sensors on these low-altitude aerial platforms enables users to acquire ultrahigh (sub–cm) spatial resolution imagery of the Earth's surface without the need for sophisticated instruments. They also overcome some of the challenges that can affect the remote collection of optical imagery, including variable atmospheric conditions (e.g., rainfall and cloud cover), via their capacity for rapid deployment [14].

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

The use of ultrahigh spatial resolution UAV-based imagery has seen increased uptake in many photogrammetric applications due to the development of computer vision techniques, such as structure from motion (SfM) and MVS reconstruction [15], [16]. The combination of MVS photogrammetry and SfM methods provides the opportunity to utilize optical UAV imagery for 3-D reconstruction, as it increases the sampling density of photogrammetry with higher spatial resolution imagery. The SfM–MVS 3-D reconstruction can now be achieved irrespective of the predefined position, orientation, and optical model information of the on-board cameras, and as a result, generally increases the accuracy of subsequent products [10], [17]. Such techniques have seen application in numerous studies but are, perhaps, most commonly employed in providing highly accurate digital surface models (DSM) for topographic surveys via the implementation of proper camera calibration and ground control points (GCP) [3], [18]–[21]. Conventionally, the imagery used for DSM reconstruction is acquired at nadir, which is similar to traditional aerial photogrammetry. James and Robson [19] suggested that, by collecting oblique images (around 20° off-nadir) in addition to nadir imagery, the accuracy of UAV-derived topographic data can be improved. However, in many natural landscapes, the survey area of interest may contain steep, overhanging, or vertical rock formations, which are unsuited for 3-D reconstruction from either traditional surveying or conventional UAV-based nadir and oblique sensor-viewing geometry.

The use of horizontal-viewing (façade) photographs in 3-D reconstruction is ubiquitous for disciplines, such as cultural heritage preservation and civil engineering [22]–[24]. Koutsoudis et al. [25] suggested that accurate 3-D modeling of artificial objects reconstructed from ultrahigh spatial resolution façade imagery can achieve similar levels of detail as other sensing methods, such as laser scanning. Man-made structures and objects are usually easier to reconstruct than natural landscapes since a range of discernible features improves the operation of computer vision algorithms [26]. In a recent application, Jalandoni et al. [27] used ground-collected façade photographs to successfully reconstruct ancient artwork on a vertical rock surface. The use of façade photographs has also proven useful for natural surface reconstruction although natural surfaces sometimes lack the distinct features required to take advantage of photograph-matching algorithms [28], [29]. James and Robson [12] used on-ground near-parallel façade photographs to reconstruct a coastal cliff section. The on-ground method was time-consuming and precluded the coverage of large areas. Therefore, acquiring façade photographs from a UAV offers an appealing alternative to increase efficiency and reduce potential data gaps [4].

The acquisition of façade or oblique UAV photographs is suitable for covering vertical walls, which are not accessible by survey personnel. Berquist et al. [30] used oblique and façade UAV imagery to build a 3-D model of an inaccessible rock cliff section, using ancient artwork on the cliff as reference points. Barlow et al. [31] acquired parallel façade photographs from a UAV to reconstruct a sea cliff to analyze rock structural discontinuities. Jaud et al. [32] assessed the impact of camera viewing angles on the reconstruction accuracy of a cliff and did not find significant differences between data sets acquired

at 20° and 40° off-nadir. In their analysis, the completeness of data reconstruction between the data sets was similar, except for areas with extruding or overhanging cliff surfaces. However, in [19], the lack of other convergent images of a coastal cliff resulted in errors related to systematic radial distortion. Such errors were likely caused by the dome effect, which often appears in SfM-produced topographic data using nadir UAV image due to the nature of bundle adjustment [18]. Circular scanning of sites, with the normal vectors of camera positions converging to a pseudopoint, is a common strategy for UAV-based collection of oblique images from near-vertical objects [33]. Martínez-Carricondo et al. [34] used the combination of nadir and oblique images to improve the accuracy of topographic data from a vertical wall in the condition of inadequate GCP distribution. Such findings indicate the potential to improve the topographic accuracy and data completeness through combining UAV imagery acquired from various viewing perspectives, especially in complex terrain containing vertical surfaces.

Here, we explore the combination of UAV-derived nadir, oblique, and façade photography to assess potential advantages for 3-D reconstruction and modeling of complex mountainous terrain with vertical rock formations. While the above studies indicate the potential benefits of façade photography [23]–[25], [31], we were unable to identify published work demonstrating UAV-based data collection procedures and processing results for integrating nadir, oblique, and façade photographs in the complex terrain. As such, the overall objective of this study was to comprehensively assess and intercompare the suitability and geometric accuracy of UAV-based image acquisition and data processing scenarios for topographic survey applications of rock formations using: 1) nadir imagery; 2) nadir and oblique imagery; 3) nadir and façade imagery; and 4) nadir, oblique, and façade imagery. The outcome of this work has implications for civil engineering, geomorphological studies, and other disciplines requiring accurate topographic surveying, especially in complex terrains where conventional approaches are restricted due to accessibility and safety concerns. The research also provides information on the benefits of integrating off-nadir and façade photographs for 3-D modeling and visualization of vertical rock surfaces, which may be required for the assessment of geological structures to evaluate rock integrity and stability in the vicinity of infrastructure.

II. MATERIALS AND METHODS

A. Study Site

The study site is located in the southwest of Saudi Arabia within the mountainous region of Aseer, covering an area of approximately 2.7 hectares. The orthometric elevation within the study area ranges from 1693–1823 m above mean sea level (see Fig. 1). The site contains many steep and vertical rock surfaces, with slopes varying from 0° to 90°. Thus, there is a requirement for repeated topographic surveys in order to estimate rock structural stability and identify potential safety issues to existing road infrastructure.

B. Data Collection and Processing

The project comprised three major data collection exercises, including the acquisition of GNSS survey data of GCPs, TLS

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TU *et al.*: COMBINING NADIR, OBLIQUE, AND FAÇADE IMAGERY ENHANCES RECONSTRUCTION OF ROCK FORMATIONS
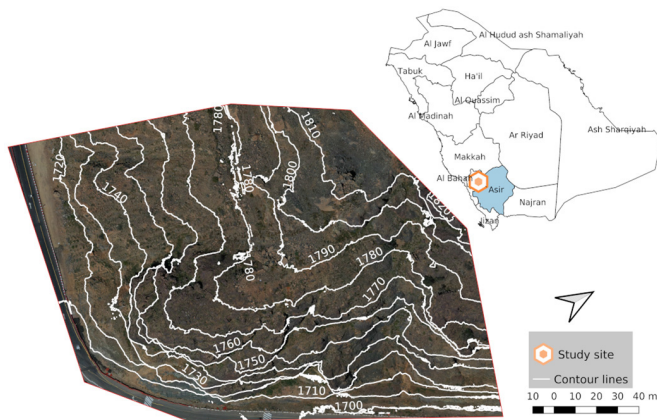
3

Fig. 1. Elevation contours derived for the study site in southern Saudi Arabia. The contour line elevation interval is 10 m. The base map is derived from a 2-cm resolution orthomosaic.
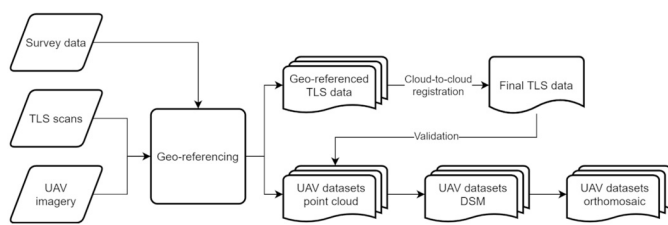

Fig. 2. Conceptual workflow of the data processing chain.
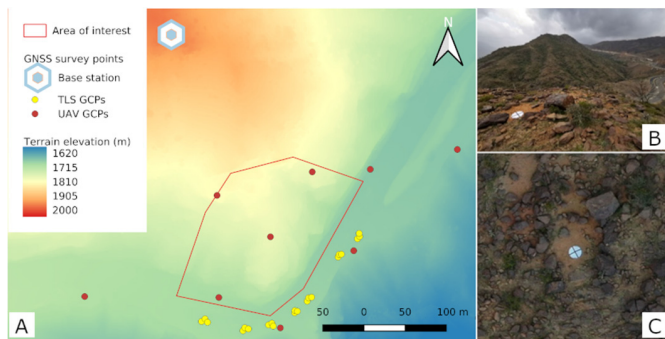

Fig. 3. (a) GNSS survey locations of GCPs used for the TLS (yellow) and UAV (red) data georeferencing, showing terrain elevation as a backdrop. (b) Example of a GCP within the study area. (c) Corresponding UAV image.

surveys of vertical rock walls, and UAV-based red–green–blue (RGB) imagery. Fig. 2 presents the workflow of the subsequent data processing, with further details of the individual processing steps provided in Section II-B.

1) GCP Deployment, Surveying, and Processing: A total of 30 GCPs were deployed for georeferencing, with 21 GCPs applied for the TLS data, and nine used for the UAV data [see Fig. 3(a)]. Portable aluminum trays marked with a black cross were used as GCPs [see Fig. 3(b) and (c)]. A Leica AS10 multi-GNSS antenna (Leica Geosystems, Heerbrugg, Switzerland) was installed as a temporary base station at a nearby high point [see Fig. 3(a)]. Static observations were collected for 60 min before starting the real-time kinematic (RTK) survey of GCPs. A pole fixed at a height of 1.8 m and integrated with a bipod was used to minimize potential misleveling of the Leica GS15 antenna rover. The raw GNSS files were downloaded and imported into the Leica Geo Office 8.4 software

(Leica Geosystems). Reference data from continuously operating reference stations (CORS) were used to accurately establish the position of the base station. The calculated position and ellipsoidal elevation of the base station were then used to postprocess the RTK data. The EGM96 geoid model was used to determine the geoid separation and applied to the ellipsoid heights for orthometric elevation determination. Results were tabulated using WGS84 as the reference datum.

2) TLS Data Acquisition and Processing: Seven TLS scans were performed using a FARO Focus X330 laser scanner (Faro Technologies, Warwickshire, U.K.) deployed on a 1.03-m tripod. Triplets of GCPs were located on the ground within 5 m of each scan location to facilitate accurate georeferencing [see Fig. 3(a)]. Each scan was performed at a distance of approximately 25 m from the rock façade. The scan resolution was 0.03° with a scan rate of 112 kpt/s. The angular area of each scan was 90° to −90° horizontally and 90° to −62.6° vertically. A photograph was taken at the end of each scan to enable rendering of the point cloud during postprocessing. The postprocessing of the TLS scans was undertaken within the Faro SCENE 5.3 software. Each scan was imported into SCENE and processed separately to generate multiple grey-scale point clouds and subsequently converted into colored point clouds based on the RGB photographs collected at the end of each scanning process. The GCPs were manually marked on each colored point cloud for georeferencing. Cloud-to-cloud registration was applied to stitch the seven separate point cloud chunks into one integrated point cloud as a reference to the rock façade. The result was exported for comparison against the UAV-derived SfM–MVS point clouds.

3) Flight Planning and UAV Image Acquisition: The UAV imaging system consisted of a gimbal-stabilized 20MP Hasselblad L1D-20c camera (Victor Hasselblad AB, Gothenburg, Sweden) and two DJI MAVIC 2 Pro quadcopters (SZ DJI Technology Co., Ltd, Shenzhen, China). The flight planning and ground controls were carried out using the universal ground control station (UgCS) application (SPH Engineering, SIA, Riga, Latvia). The flight campaign consisted of three nadir-, three oblique-, and six façade-viewing flights. While all nadir-viewing flights and five façade scans were preplanned for autonomous flight missions, one façade scan and three oblique scans of the rock wall were flown manually due to safety and potential collision concerns. We collected oblique and façade scan images in the mornings and afternoons from sunlit rock walls to reduce shadowing effects, whereas nadir-viewing images were collected within two hours of solar noon. Table I shows the parameters of the different flight configurations. The distance between the sensor and the terrain surface was not consistent due to the extreme terrain variation. The three nadir-viewing flights were flown at various heights (1850, 1860, and 1910 m) above mean sea level to ensure safe clearance of the terrain. It should be noted that the camera viewing elevation of façade flights was set to 4° to avoid the influence of the propellers.

4) UAV Image Processing: The collected UAV images were organized into four data sets: 1) nadir imagery (1287 photographs); 2) nadir and oblique imagery (1932 photographs); 3) nadir and façade imagery (3297 photographs); and 4) nadir,

TABLE I
Flight Parameters of the UAV Campaigns for Nadir, Façade, and Oblique Photography

Parameters	Nadir	Façade scans	Oblique scans (manual)
Number of flights	3	6	3
Number of photos	1287	2010	645
Height above ground (m)	67-197	6-80	20-120
Distance to rock (m)	67-197	~20-40	~20-40
Speed (m/s)	3	2	Manual
Side distance (m)	10	3.5	~20
Forward overlap (%)	> 90	> 77	> 75
Sidelap (%)	>88	~86-93	Variable
Camera viewing elevation (°)	90	4	~30

oblique, and façade imagery (3942 photographs). The four data sets were imported into Agisoft Metashape Pro 1.6.2 (Agisoft LLC, Saint Petersburg, Russia) for processing. The image processing workflow was the same for the four data sets. The processing workflow references the USGS guidelines [35]: particularly, the error reduction methods to ensure data accuracy. The coordinates of the photographs were converted from a WGS84 geographic coordinate system to a UTM projected coordinate system. The predefined camera model was checked before the initial photograph alignment. Photograph alignment was performed at the original scale, with a tie point limit of 4000 and a key point limit of 40 000 points. Once the photograph-alignment was completed, error reduction was performed based on reconstruction uncertainty and projection accuracy to remove those tie points with an accuracy below the USGS-suggested thresholds without omitting any aligned photographs. Values of ten for reconstruction uncertainty and two for projection accuracy were defined as the thresholds for error reduction at the initial phase.

The GNSS-derived GCP coordinates were subsequently imported and manually registered to the GCPs visible in the UAV images, followed by another error reduction based on the reprojection error. The remaining tie points with an accuracy below the USGS-suggested threshold of 0.3 pixels were removed without omitting aligned photographs. This enabled the completion of the SfM-derived sparse point cloud. The root-mean-square errors (RMSEs) for data sets 1 (nadir), 2 (nadir and oblique), 3 (nadir and façade), and 4 (nadir, oblique, and façade) based on the GCPs were 1.55, 0.89, 1.84, and 1.37 cm, and the projection errors were 0.4, 0.49, 0.56, and 0.6 pixel, respectively. MVS dense clouds were then generated using the original scale (ultrahigh-resolution) and aggressive filtering for all data sets. Any visually identified noise in the dense point clouds was subsequently removed. Finally, the results were exported for comparison with the TLS-derived point cloud data. The point density of the data sets 1, 2, 3, and 4 was 0.13 points/cm^2, 0.16, 0.19, and 0.2 points/cm^2, respectively. The produced SfM–MVS point cloud of each data set was then used to generate a DSM with a Poisson surface reconstruction method [36], followed by the generation

of an orthomosaic based on the DSM and the mosaic blending mode.

5) Accuracy Assessment: The accuracy of the UAV-derived point clouds was evaluated against the georeferenced TLS-derived point cloud for the rock façade using the CloudCompare 2.12 alpha version [37]. The first step was to crop the data with the polygon of the study site to ensure consistent data volumes. Because the seven TLS scans were merged as one integrated cloud, geometric errors within each point cloud may be propagated during the process [38]. Therefore, the iterative closest point (ICP) fine registration algorithm [39] was used to calculate the rigid body transformation matrix to fit the TLS data to the UAV data sets (1). ICP fine registration assumes that the two comparing point clouds are roughly aligned and the shapes in the overlapping area are the same and only transform points at a fine scale. The rigid body transformation matrix has the same unit as the multiplying coordinate vectors. The multiplication of such a matrix and the multiplying coordinate vectors (easting, northing, height, and scale in a single column) results in the transformed coordinate vectors, which maintains the original unit (m). The transformation matrix was eventually applied to the TLS data before the comparison with each UAV-derived point cloud to minimize the geometric error between the TLS data and each UAV data set.

Each finely registered TLS data set was then compared with the corresponding UAV-derived point cloud data. Because the TLS data had smaller data coverage than the UAV-derived point cloud data, it was set as the comparing data rather than the reference data in CloudCompare due to the restriction of the comparison algorithm employed by the software. Nevertheless, the magnitude of the resulting cloud-to-cloud alignment errors was the same regardless of which data set was assigned as comparing and reference data, and hence, the results provided an indication of the errors of the UAV data sets when evaluated against the TLS data. For every comparing point from the TLS data, a local surface model was calculated using a least-squares method with the nearest six points from the corresponding UAV-derived point cloud data. The absolute Euclidean distance between each point from the TLS point clouds and the corresponding local least-squares plane was calculated to determine the geometric accuracy of the rock façade within the four UAV-derived point clouds based on the nadir; nadir and oblique; nadir and façade; and nadir, oblique, and façade imagery, respectively.

Subsequently, the UAV-derived point cloud, which had the lowest geometric error, was selected as the reference point cloud for relative comparison with the three remaining UAV-derived point clouds. Such comparisons further assessed the point cloud accuracy for those parts of the study site that was not covered in the TLS-based point clouds. The comparison method for the UAV-derived data was the same as that used for the TLS data, except no ICP fine registration was applied before the comparison. The absence of ICP on UAV-derived data is because all the UAV-derived data were georeferenced based on the same GCPs, and no further geometric manipulation of the UAV-derived point clouds was applied after GCP registration. In addition, the raster DSM data of each UAV data set were used to subtract the DSM
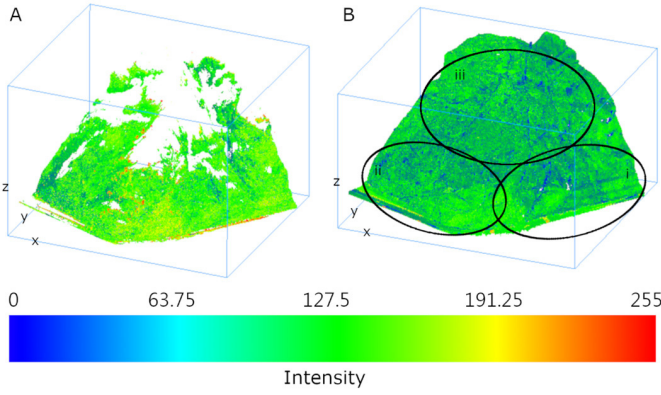
Fig. 4. Data coverage of (a) TLS point cloud and (b) point cloud derived from nadir UAV images. The point cloud colors default to the 8-bit intensity ranged from 0 to 255 and represent the reflected energy strength of the TLS data and the grey scale of point colors of the UAV-derived data, respectively. The three regions outlined in (b) represent: 1) east-facing façade; 2) south-facing façade; and 3) top of the rock formation.

of the selected UAV reference data set to more intuitively demonstrate the elevation differences, a process that was performed using QGIS [40]. Finally, the orthomosaic of each UAV data set was visually compared with the orthomosaic of the same UAV reference data set to check for significant spatial offsets.

III. RESULTS

A. Point Cloud Comparison Between the TLS and UAV-Derived Data

The collection of TLS data enabled the coverage of the lower sections of the rock formation up to a height of approximately 70 m [see Fig. 4(a)]. The scanned area was roughly divided into three parts, as shown in Fig. 4(b), with areas representing the east-facing façade, south-facing façade, and top of the rock formation. These defined regions are used to inform the analysis presented in Sections III and IV. The coverage of the TLS data was concentrated on the east- and south-facing façade sections due to the reduced line of sight. This is particularly obvious in the central area (where sections i and ii join) and at the top of the rock formation (section iii) due to the near-vertical rock wall obstructing the view of the area above. In the lower sections, some spots were also omitted in the TLS point cloud due to obstructing rocks and the fact that only seven scans (i.e., points of viewing) were used to create the point cloud. In contrast, the UAV imagery consists of thousands of viewing points, providing a more complete point cloud of the area of interest [see Fig. 4(b)].

The rigid body transformation matrices, which were applied to the TLS data before each comparison, are presented

in the following. Transformations were necessary to minimize the influence of the geometric propagation errors caused by the merging of multiple TLS scans. They also ensure that the derived cloud-to-cloud errors between the TLS data and UAV point clouds could be meaningfully interpreted. The meaning of each element in the matrix can be found in (1), as shown at the bottom of the page.

The rigid body transformation matrix to fit the TLS point cloud to UAV data set 1 (nadir images) was

$$\begin{pmatrix} 1 & 0 & 0.001 & -1.549 \\ 0 & 1 & 0.003 & -4.674 \\ -0.001 & -0.003 & 1 & -0.076 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The transformation matrix to fit the TLS point cloud to UAV data set 2 (nadir and oblique images) was

$$\begin{pmatrix} 1 & 0.001 & -0.002 & 3.521 \\ -0.001 & 1 & 0.009 & -15.011 \\ 0.002 & -0.009 & 1 & -0.409 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The transformation matrix to fit the TLS point cloud to UAV data set 3 (nadir and façade images) was

$$\begin{pmatrix} 1 & 0.001 & -0.002 & 2.761 \\ -0.001 & 1 & 0 & -0.718 \\ 0.002 & 0 & 1 & -0.101 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Finally, the transformation matrix to fit the TLS point cloud to UAV data set 4 (all images) was

$$\begin{pmatrix} 1 & -0.001 & -0.002 & 2.62 \\ 0.001 & 1 & 0.003 & -5.54 \\ 0.002 & -0.003 & 1 & 0.048 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The matrices show that the horizontal offsets (the combination of the X shift along the x-axis and the Y shift along the y-axis) were two orders larger than the vertical offsets (the Z shift along the z-axis) in UAV data sets 1 and 4 and at least one order larger in UAV data sets 2 and 3. This may be due to the TLS data being concentrated on the façade parts of the rock formation and hence causing a lack of discernible features on the horizontal plane. The lack of horizontal features increased the geometric propagation errors in the horizontal directions during the cloud-to-cloud alignment process when merging the TLS scans. Both the Y shift along the y-axis (-15.011 m) and the Z shift along the z-axis (-0.409 m) were an order larger in the transformation matrix (which fits the TLS data to UAV data set 2) than the other transformation matrices. This indicates that a larger transformation was applied to the TLS

$$\begin{pmatrix} \cos\varphi\cos\psi + \sin\varphi\sin\theta\sin\psi & \sin\varphi\cos\psi - \cos\varphi\sin\theta\sin\psi & \cos\theta\sin\psi & X \\ \quad -\sin\varphi\cos\theta & \cos\varphi\cos\theta & \sin\theta & Y \\ \sin\varphi\sin\theta\cos\psi - \cos\varphi\sin\psi & -\cos\varphi\sin\theta\cos\psi - \sin\varphi\sin\psi & \cos\theta\cos\psi & Z \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (1)$$

where θ, ψ, and φ are the rotations around the x-, y-, and z-axes, and X, Y, and Z are the shifts along the x-, y-, and z-axes, respectively.

data to reduce the Euclidean distances to the point cloud of UAV data set 2. The larger transformation also means that larger relative errors in position exist between the TLS data and the point cloud of UAV data set 2. Therefore, we can expect that UAV data set 2 would have a larger difference in point locations compared with other UAV-derived data with lower transformation coefficients.

The mean absolute point distance (point number ≈ 81.54 million) between the rigid-body-transformed TLS data and UAV data set 1 was 5.67 cm, with an 8.86-cm standard deviation. The mean absolute distance against UAV data set 2 was 5.74 cm, with a standard deviation of 6.80 cm. Although the mean absolute point distance of these two data sets was similar, UAV data set 2 was still considered to have a poorer geometric accuracy because the magnitude of the shift distance (the fourth column) in the transformation matrix was an order larger than the shift distance of UAV data set 1. The mean absolute distance against the point cloud of UAV data set 3 was the largest among the four data sets, with a mean absolute distance of 9.46 cm and a 10.48-cm standard deviation. Meanwhile, the difference between the rigid-body-transformed TLS data and UAV data set 4 was the smallest, with a mean absolute distance of 3.71 cm and a 5.43-cm standard deviation. Since the transformation matrices to fit the TLS data to UAV data sets 1, 3, and 4 were of similar magnitude for the data transformation, we can conclude that the transformed TLS data had the smallest geometric errors against UAV data set 4, which also means that the geometric errors between UAV data set 4 and the TLS data were the smallest among the four UAV data sets. The improvement in geometric accuracy of UAV data set 4 was found to be 34.57% compared with UAV data set 1.

The percentage of points that had a distance error of less than 50 cm was 99.56% (data set 1), 99.92% (data set 2), 99.47% (data set 3), and 99.96% (data set 4) for the respective point clouds. From Fig. 5, it can be seen that not only did data set 4 have the smallest distance error but it also had the most homogeneous error distribution within the area covered by the data [see Fig. 5(d)], indicating the similarity between the TLS and data set 4 point clouds. Among data sets 1, 2, and 3, data set 1 had smaller distance errors (blue) distributed more homogeneously than the other data sets, except for areas where protruding rocks appeared within the near-vertical façades. This emphasizes that the nadir-viewing imagery precluded accurate reconstruction of vertical objects. On the other hand, larger errors (green and red) were distributed more evenly on the south-facing façade [see Fig. 5(b)] in both data sets 2 and 3, whereas the errors with the same magnitude in data set 3 were mostly distributed at the bottom of the south-facing façade [see Fig. 5(c)]. While some of the TLS data reached the top of the rock formation, UAV data set 4 still maintained lower errors than the other three data sets [see Fig. 5(d)]. Specifically, for vertical and overhanging rock sections, UAV data set 4 performed the best, which is illustrated in Fig. 6, where the point distance between the TLS point cloud and the point cloud of UAV data set 4 was visibly lower (<20 cm) than the point distance between the TLS and UAV data set 1 point clouds (up to 50 cm for overhanging rocks). Based on these
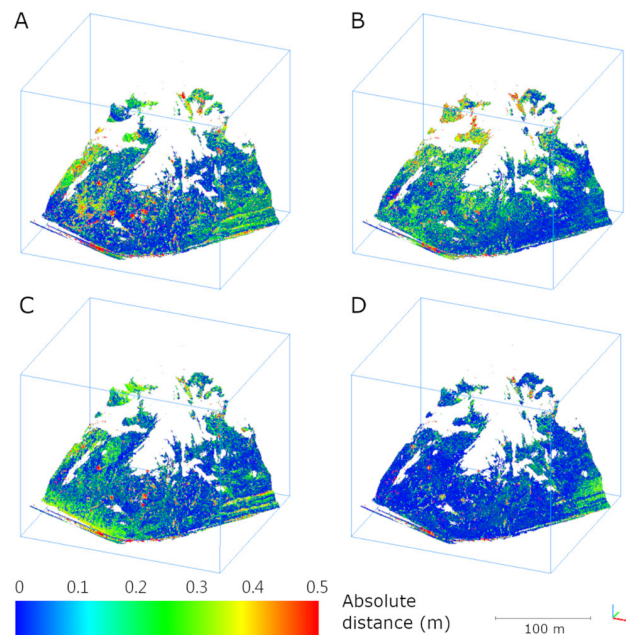

Fig. 5. Point distances between the rigid-body-transformed TLS data and (a) UAV data sets 1 (nadir), (b) data set 2 (nadir and oblique), (c) data set 3 (nadir and façade), and (d) data set 4 (nadir, oblique, and façade). The color ramp from blue to red represents the cloud-to-cloud absolute distance.
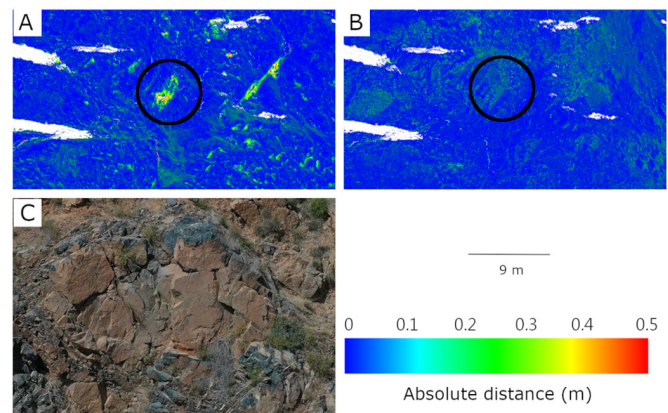

Fig. 6. Zoomed-in view of a small section of the south-facing façade with an overhanging rock surface (black circles), showing point distances between (a) TLS data and UAV data set 1 (nadir) and (b) TLS data and UAV data set 4 (nadir, oblique, and façade). (c) Specific overhanging rock surface from the raw façade image (scale bar not applicable). White areas in (a) and (b) represent data gaps due to the TLS line-of-sight viewing limitation.

observations, it can be concluded that UAV data set 4, which comprised the nadir, oblique, and façade imagery, had the lowest errors when evaluated against the TLS data. As such, it was selected for use as reference data against which the accuracy of the other UAV data sets could be assessed.

B. Point Cloud Comparison Between the UAV Data Sets

Unlike the comparison between the TLS data and UAV data sets, no rigid body transformation was required for comparing the UAV-derived point clouds. Therefore, it was expected that the resulting distance errors between the UAV data sets would be larger than the errors observed in the previous section. The mean absolute distance between data sets 1 and 4 [see Fig. 7(a)] was 5.69 cm, with a standard deviation of 8.8 cm

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TU *et al.*: COMBINING NADIR, OBLIQUE, AND FAÇADE IMAGERY ENHANCES RECONSTRUCTION OF ROCK FORMATIONS
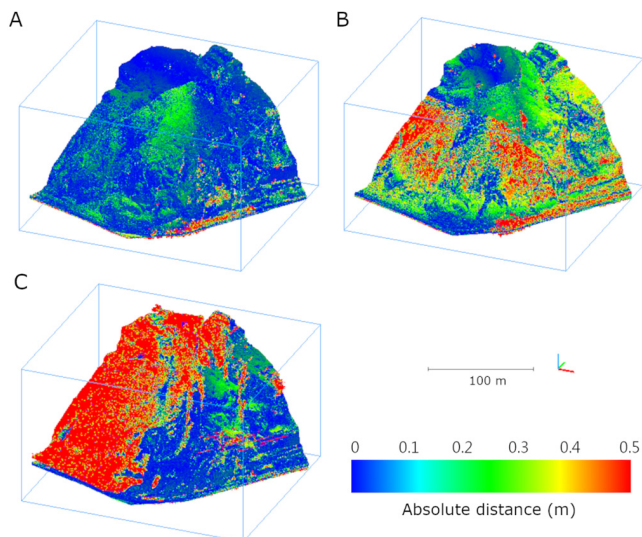
7

Fig. 7. Point distances between (a) UAV data sets 1 (nadir) and 4 (nadir, oblique, and façade), (b) data sets 2 (nadir and oblique) and 4, and (c) data set 3 (nadir and façade) and 4, with data set 4 used as the reference data set. The color ramp from blue to red represents the cloud-to-cloud absolute distance.

TABLE II

STATISTICS OF THE POINT-TO-POINT DISTANCES BETWEEN UAV DATA SET 4 AND THE OTHER THREE UAV DATA SETS, INCLUDING MEAN ABSOLUTE DISTANCE (μ), STANDARD DEVIATION (σ), AND THE PERCENTAGE OF DISTANCE ERRORS <50 CM

dataset	Euclidean distance (μ/σ)	Horizontal distance (μ/σ)	Vertical distance (μ/σ)	Point errors < 50 cm
1	5.69 cm / 8.8 cm	3.47 cm / 4.46 cm	3.65 cm / 7.27 cm	99.41%
2	18.05 cm / 30.75 cm	10.60 cm / 19.01 cm	8.02 cm / 16.92 cm	90.38%
3	36.75 cm / 72.32 cm	24.19 cm / 53.36 cm	17.82 cm / 44.40 cm	80.19%

(point number = 169 244 054 points). Comparing data sets 2 and 4 [see Fig. 7(b)], the mean absolute distance increased to 18.05 cm, while the standard deviation was 30.75 cm (point number = 517 919 500 points). The errors between data sets 3 and 4 [see Fig. 7(c)] were the highest among the three comparisons, with a mean absolute distance of 36.75 cm and a standard deviation of 72.32 cm (point number = 588 152 698 points). A total of 99.41% of points in data set 1 had errors < 50 cm, while only 90.38% and 80.19% of points in data sets 2 and 3 achieved the same accuracy threshold, respectively (see Table II). If the absolute distance was interpreted into absolute horizontal and vertical distances, UAV data set 1 had a smaller horizontal mean distance of 3.47 cm with a standard deviation of 4.46 cm compared with the vertical mean distance of 3.65 cm with a standard deviation of 7.27 cm. UAV data sets 2 and 3 both had larger absolute point distances in horizontal directions of 10.60 and 24.19 cm and vertical directions of 8.02 and 17.82 cm, respectively (see Table II).

From Fig. 7, we can tell that the absolute mean distance between data sets 1 and 4 was generally less than 10 cm, while the geometric alignment errors were much larger between data sets 2 and 4, as well as 3 and 4. The comparison
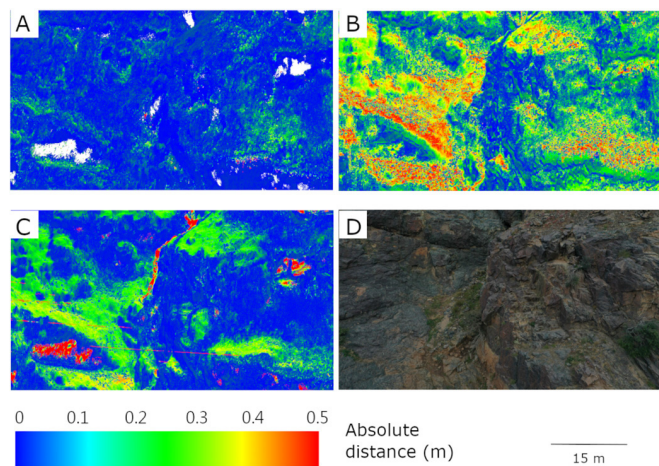

Fig. 8. Zoomed-in view of the east-facing façade of the point distances between (a) UAV data sets 1 (nadir) and 4 (nadir, oblique, and façade), (b) data sets 2 (nadir and oblique) and 4, and (c) data set 3 (nadir and façade) and 4, with data set 4 used as the reference data set. A raw image from one of the façade flights (d) covering the approximate area is also presented to assist in visualizing the morphology of the terrain.

between data sets 2 and 4 exhibited the largest point distances along both the east- and south-facing façade sections but with smaller geometric offsets occurring in the top sections of the rock formation [see Fig. 7(b)]. The comparison of data sets 3 and 4 showed the largest point distances along the south-facing façade and the top of the rock formation, whereas the east-facing section of the façade only displayed minor offsets between points generally of less than 10 cm [see Fig. 7(c)]. This phenomenon may be caused by differences in the rock surface-to-camera viewing angle between the two façade sections. The east-facing façade had steeper terrain than the south-facing façade (see Fig. 1), causing the south-facing façade to have a larger angle between the viewing direction of the façade photographs and the normal vector of the façade surface. Smaller viewing-surface geometry causes smaller ground sampling distance (GSD) variation, which provides higher topographic survey quality [29].

Fig. 8 shows a zoomed-in area of the point distances focused on the east-facing façade, where the steepest slopes were located (see Fig. 1). As shown in Fig. 8(a), there were gaps in data set 1 (i.e., comprised of only nadir images), indicating a limitation in reconstructing the terrain in areas with large terrain fluctuations, or where overhanging rocks were present. Although adding either oblique images [see Fig. 8(b)] or façade images [see Fig. 8(c)] achieved near-perfect data completeness, the accuracy was less than satisfactory compared with the reference data, which integrated the nadir, oblique, and façade imagery.

To further explore the data, we calculated the least-squares linear regression between the absolute point distance and the terrain slope. The least-squares residual of each comparison was 9 cm between UAV data sets 1 and 4, 22 cm between data sets 2 and 4, and 32 cm between data sets 3 and 4. Overall, there was no linear relationship established between point distance and terrain slope identified for any of the data set comparisons. Extreme point distance values (mean value plus twice the standard deviation) occurred more frequently
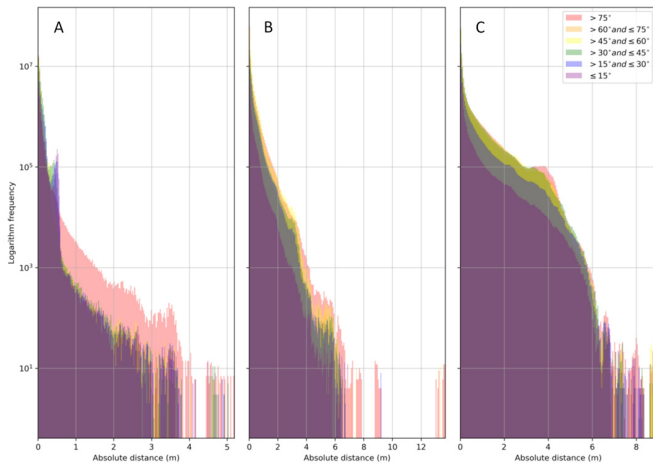
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

8

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

Fig. 9. Histograms of point distances within different ranges of terrain slopes of (a) UAV data set 1, (b) data set 2, and (c) data set 3. The *y*-axis represents the frequency of points is displayed on a logarithm scale.
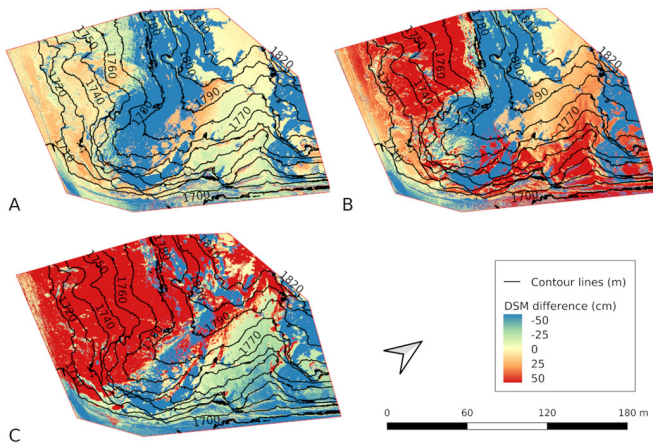

Fig. 10. DSM elevation differences when subtracting the reference UAV data set 4 (nadir, oblique, and façade) from UAV data sets (a) 1 (nadir), (b) 2 (nadir and oblique), and (c) 3 (nadir and façade).

(10.06% for data set 1, 10.60% for data set 2, and 9.68% for data set 3) when the slope was between 85° and 90°. Fig. 9 shows the histograms of point distance in various ranges of terrain slope. The extreme point distances occurred more frequently on steep slopes, especially for the comparison of UAV data sets 2 and 4, where the maximum values of the extreme point distance increased up to 14 m [see Fig. 9(b)]. The occurrence frequency of points with large distances relative to those in data set 4 seems to increase significantly for data set 1 when the slope was higher than 75° [see Fig. 9(a)]. However, the frequency of points with large distances relative to those in data set 4 for data sets 2 and 3 was not much different irrespective of terrain slope [see Fig. 9(b) and (c)].

C. DSM and Orthomosaic Comparisons Between UAV Data Sets

The DSMs produced from the point cloud of UAV data sets 1, 2, and 3 were also compared with the DSM produced from data set 4 by subtracting the DSM raster images from one another (see Fig. 10). Similar DSM elevation differences of more than −50 cm (blue) occurred at the top of the rock
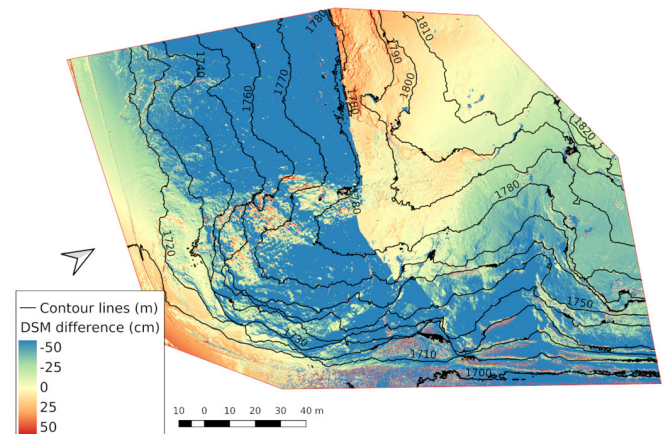

Fig. 11. DSM elevation difference between UAV data sets 1 and 2.

formation (above the 1730-m contour line on the east part of the east-facing façade) across the three comparisons. The same response was observed above the 1760-m contour line at the south-facing façade in the comparisons of data sets 1 and 2 to data set 4 [see Fig. 10(a) and (b)]. Such consistent differences across all comparisons indicated that the DSM errors on the top of the rock formation may be due to the reference DSM of UAV data set 4 itself.

Fig. 11 shows the DSM difference between UAV data sets 1 and 2 and indicates that the elevation differences on the top of the rock formation were significantly smaller, particularly compared with the DSM difference between either UAV data sets 1 and 4 or between UAV data sets 2 and 4 (see Fig. 10). As UAV data set 1 only comprised the nadir imagery and had the smallest projection error (0.4 pixels), the geometric accuracy of surfaces where the nadir imagery could easily observe should be higher, considering that the RMSE of tie points on GCPs was similar between UAV data sets 1 and 4. Hence, it may indicate that UAV data set 4 did not produce an accurate DSM for the top section of the rock formation, regardless of the smaller errors observed in the point cloud data of the façade sections [see Fig. 5(d)]. Overall, the DSM of UAV data set 1 was most similar to the reference DSM of UAV data set 4. Surprisingly, the DSM difference between UAV data sets 1 and 4 on the top section of the rock formation was higher (>50 cm) than the point distances shown in Fig. 7(a), which were around 30 cm, with only 0.59% of points in data set 1, exhibiting a distance of >50 cm compared with UAV data set 4. This may be because the surface reconstruction algorithm used by the software was affected by the small percentage of highly offset points, thus accumulated errors when producing the DSM.

All orthomosaics displayed a systematic issue of duplicating objects, as well as some minor offsets between data sets. Fig. 12 demonstrates a systematic issue causing duplication of objects at the lower part of the south-facing façade and on adjacent flat ground. This issue appeared in the same area in all the UAV data sets and required manual seamline correction.

Fig. 13 shows minor offsets (range from 5 cm to 1 m) of specific objects between UAV data set 4 and the other

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TU *et al.*: COMBINING NADIR, OBLIQUE, AND FAÇADE IMAGERY ENHANCES RECONSTRUCTION OF ROCK FORMATIONS 9
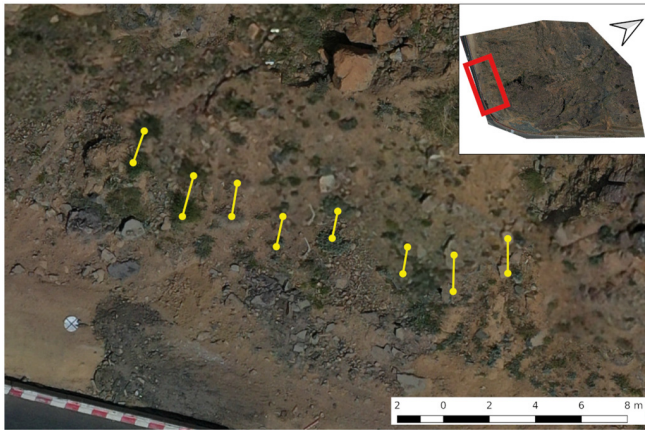

Fig. 12. Duplicated object effect caused during the generation of the orthomosaics within the area of the red rectangle. The ball caps at the ends of each yellow line indicate some rocks and small bushes appearing twice.
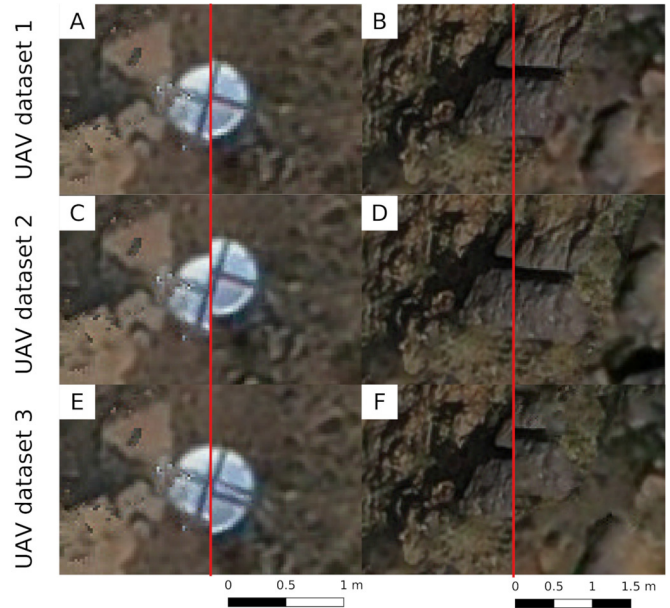

Fig. 13. Swipe comparisons between each UAV data set (1, 2, and 3) and UAV data set 4 of selected locations. The images at the left-hand side of the swipe lines show a part of specific objects in the mosaicked product of UAV data set 4, while the image at the right-hand side of the swipe lines represent the other part of the same objects from UAV data sets 1, 2, and 3. (a), (c), and (e) offset of the same GCP. (b), (d), and (f) Offset of the same rock between UAV data set 4 and the respective data sets.

UAV data sets. From the analysis in Section III-B, the horizontal distance between UAV data sets 1 and 4 was found to be smaller than the vertical distance. Yet, the opposite was observed between UAV data sets 2 and 4 and between UAV data sets 3 and 4. Only small horizontal distances were observed between the orthomosaics in the areas around the GCPs, exhibiting offsets of 5 cm [see Fig. 13(a) and (e)] to 10 cm [see Fig. 13(c)]. However, larger offsets were found in areas without GCPs [e.g., 20 cm in Fig. 13(b) and 50 cm in Fig. 13(d) and (f)]. The horizontal offsets found in the orthomosaics were larger than the horizontal distance in the point cloud alignment analysis, where the mean horizontal distance was 3.47 and 10.60 cm for UAV data sets 1 and 2, respectively. This may be due to the potential error accumulation during the production of DSMs, as the orthomosaic image products were built based on the DSMs rather than rasterizing the point cloud on a projected coordinate plane directly. Besides, a traditional mosaicking algorithm is likely to be more suitable for near-parallel image data. It is possible that the issues observed in the orthomosaics were due to stitching problems, where inappropriate perspectives of the images were chosen by the stitching algorithm (which can be fixed by manual seamline editing).

IV. DISCUSSION

A. Influence of Photograph Composition on SfM in Complex Terrains

The inclusion of both oblique and façade photographs, in addition to traditional nadir images, generally improved the geometric accuracy of the point cloud data retrieved from complex terrain in the surveyed rock formation. Although the point positions between data set 1 (produced from only nadir images) and data set 4 (produced from the combination of nadir, oblique, and façade images) were similar for the façade sections (i.e., an overall least-squares residual of 9 cm), the additional oblique and façade photographs improved the geometric accuracy when evaluated against the TLS data, especially in areas with large terrain fluctuations and where overhanging rocks were presented [see Figs. 6 and 8(a)]. However, when adding either the oblique or façade images

to the nadir photographs (data sets 2 and 3), the geometric accuracy worsened, especially in the south-facing façade area. Such a decrease in accuracy might be due to systematic radial topographic distortion caused by the bundle adjustment method used by the processing software. James and Robson [19] demonstrated that the bundle adjustment method can overestimate the surface terrain at the center of the site while underestimating the surface terrain at the site perimeters. Such radial distortion can be reduced by deploying GCPs at suitable locations [10]. However, there were no GCPs on the slope and façade to constrain the locations during the process of bundle adjustment, which was likely the main reason behind such high errors on the sloping and façade areas in data sets 2 and 3. To prevent such distortions with limited use of GCPs, previous studies suggested using predefined camera models and an accurate onboard GNSS system to reduce the errors in topographic data [10], [41]. The lack of transitioning images in data set 3 (i.e., images collected at a viewing angle in between the nadir and façade viewing angles) was likely making the geometric accuracy worse. The bundle adjustment method was designed to be more suitable for matching convergent images that have transitioning viewing angle differences [33]. A distinct change of the camera viewing angle of UAV photographs may affect the accuracy of the generated topographic data. Based on the results presented here, it is advised to use the combination of nadir, oblique, and façade images to prevent high topographic errors—-especially when façade information is needed, e.g., for producing high-quality texture layers for the rendering of 3-D models suited for visual assessment or classification of features on steep or vertical rock walls, or if the geometric accuracy of façades is important, e.g., for slope stability assessment. Similar findings were also

suggested in the survey of a dam when GCP distribution was inadequate [34].

B. Potential Influence of Surface-Viewing Geometry

The large angle between the viewing direction of photographs and the normal vectors of the terrain surface, especially the viewing directions of the façade photographs against the top section of the rock formation (approximately 70°), maybe another reason for the large geometric errors in the top section of the rock formation in UAV data set 4 (see Fig. 10). Furthermore, the distance of the UAV to the top section when collecting the façade photographs was >200 m. The large distance and surface-viewing geometry increase the GSD variation of the photographs and result in the poor-quality photograph, which decreases the geometric accuracy of SfM–MVS produced topographic data [29]. The large difference in point distances between UAV data sets 3 and 4 on the two different façade sections may also be due to the large viewing-surface geometry. The east-facing façade had a steeper terrain than the south-facing façade, with smaller surface-viewing geometric variation between the façade photographs and the terrain surface occurring at the east-facing façade, and thus, smaller point distances were observed compared with the point distances of the south-facing façade [see Figs. 7(c) and 10]. Two potential solutions exist. The first solution is to iteratively execute the error reduction steps until reaching the accuracy criterion, which might result in the omission of some photographs during the process. While omitting photographs during the error reduction steps could have significantly reduced errors and DSM elevation differences at the top section of the rock formation, removing too many photographs would introduce gaps in the produced topographic data because of insufficient image overlap. Hence, users must test how many iterations of error reduction are needed before compromising the accuracy at the expense of data completeness. The second solution is to restrict the façade photograph acquisition to capture only the façade parts of the terrain, as suggested in [32]. In this case, the 3-D reconstruction at the top of the rock formation would only consider nadir and oblique images. The smaller the GSD variation in the photographs, the higher the quality of the topographic data achieved by the MVS method [42].

C. Influence of Redundant Information From Input Photographs

When processing the UAV-derived point clouds, visible noise was removed to improve the quality of the data. In this study, one of the most significant noise contributors was introduced by the additional façade photographs and their inclusion of sky color. From this perspective, capturing only the façade parts of the terrain can prevent not only those errors caused by the large angle between photograph-viewing direction and the normal vector of the terrain surface but also the potential influence of sky color from façade photographs. Although most of the sky-introduced noise can be removed during the error reduction phases and the point cloud densification step (with the aggressive filter applied), the color of the sky was still rendered on the MVS dense cloud result, which

requires manual removal. If the sky above the horizon of the upper sections of façade scans is included in the photographs, it is suggested to mask out unnecessary information (i.e., the section of the sky above the horizon) in the images before the photograph-alignment step in Metashape Pro [43]. Such a process could be semiautomated by applying a color filter to mask the sky in a large number of photographs [44]. Also, the difference in illumination conditions between the different flights and shadows may also affect the accuracy of SfM–MVS–generated point clouds [45].

D. Impact of Photograph Input Amount on Processing Efficiency

The ability to collect façade photographs to produce topographic data is mostly limited to local-scale analysis. Despite the potential accuracy improvements from the inclusion of oblique and façade photographs, the most significant drawback is the additional processing time due to the increased number of photographs. The machine that ran the image processing comprised of 56 virtual central processing units in 2.4-GHz Broadwell Intel Cores, four NVIDIA Tesla M10 graphics processing units, and 480-GB random access memory on an Ubuntu 16.04 system. It took 80 h (data set 1; 1287 photographs), 101 h (data set 2; 1932 photographs), 199 h (data set 3; 3297 photographs), and 290 h (data set 4; 3942 photographs) to complete the photograph matching, photograph alignment, depth map, and dense cloud generation. The processing time versus the input photograph number grew exponentially, similar to the findings of [46]. To compensate for this issue, the alternative is to process the MVS point cloud densification in multiple smaller chunks separately after the initial SfM tie points are created and the georeferencing and error reduction steps are completed. By splitting the original data into 3×3 chunks, the processing time of data set 4 was reduced by approximately two-thirds in the experiment. Despite these constraints, the incorporation of façade photographs can provide significant extra value to topographic data to accommodate various UAV applications [47].

E. Impacts of Photograph Inputs on Derived Raster Products

A factor that requires further research is the influence of the subsequent survey products, especially the production of the orthomosaic. In previous studies where façade photographs were included, the final products were usually either a point cloud, 3-D model, or an orthomosaic in a local planar coordinate system, rather than a real-world projected coordinate system [32], [47], [48]. Conventionally, the orthomosaic produced for topographic survey applications would be composed mainly of nadir aerial images with the normal vectors approximately perpendicular to the ground surface [43], [49]. Henriques *et al.* [50] suggested that factors such as the lack of features, the small coverage of each photograph, and areas full of similar elements can significantly reduce the accuracy of both the produced DSM and orthomosaic. In our case, each photograph covered a relatively small area due to the close range of the UAV to the rock surfaces and had similar repeating elements (e.g., rocks) appearing frequently

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TU *et al.*: COMBINING NADIR, OBLIQUE, AND FAÇADE IMAGERY ENHANCES RECONSTRUCTION OF ROCK FORMATIONS 11

within the study site. Therefore, it is likely that besides the accumulated errors found in the DSMs, the characteristic of repeating elements, as well as small spatial coverage of each photograph, caused difficulties for the image stitching algorithm to select the optimal orthorectified images from appropriate perspectives to generate an accurate orthomosaic. Tu *et al.* [51] suggested that flying at a higher altitude with a higher resolution camera to acquire a larger image cover area (as well as smaller GSD) could potentially increase the topographic data quality. However, photographs with ultrahigh spatial resolution collected at a higher altitude may cause higher GSD variation within scenes in such a scenario, as they may capture both local minimum and maximum elevations in a scene with fluctuating terrain. High variation in GSD could introduce propagation errors caused by reconstruction uncertainty [42]. Therefore, it is not clear if this principle is applicable in complex terrain. Further research is required to identify the causality between flight parameters and effects on derived UAV-based products (i.e., point cloud, DSM, and orthomosaic) in a common processing workflow to develop guidelines on the collection and processing of UAV photographs for topographic survey purposes in complex terrain.

V. CONCLUSION

An assessment of the impact on geometric accuracy of UAV data collected in the form of nadir, oblique, and horizontally viewing imageries in a complex mountainous terrain was undertaken. The inclusion of oblique and façade imagery, in addition to nadir photographs, improved the geometric accuracy of SfM–MVS point cloud data when assessed against TLS data of near-vertical rock façades. The nadir viewing UAV photographs produced the second-highest geometric accuracy, while the lowest geometric accuracies were obtained when processing nadir and façade photographs together in the same assessment against the TLS data. The poor geometric accuracy when combining nadir and façade photographs was related to the lack of transitioning images and the large angle between the photograph-viewing direction and the normal vector of the terrain surface. Users should avoid acquiring images with a large angle between the viewing direction of the photograph and the corresponding normal vector of the terrain surface and a distinct change of camera viewing direction, which may reduce the accuracy of the produced topographic data. Despite issues such as increased processing time and potentially more noise caused by unwanted information, such as sky color in the point cloud data (which can be addressed with alternative workflows), the accuracy improvement when integrating nadir, oblique, and façade photographs and the higher resolution façade details may provide additional benefits and application value to the produced topographic data. Applications that may benefit from the additional spatial resolution of rock walls and higher geometric accuracy include geohazard investigation, studies assessing rock integrity and stability, and engineering and construction designs relying on detailed information of geological structures, such as rock cracks, joints, faults, shears, and bedding planes. Future studies should focus on establishing an optimal protocol for the acquisition and processing of UAV imagery of complex terrain, particularly those

comprising steep or near-vertical terrain elements, so that produced products can meet the requirements of topographic survey applications.

REFERENCES

[1] P. Baldi, M. Coltelli, M. Fabris, M. Marsella, and P. Tommasi, "High precision photogrammetry for monitoring the evolution of the NW flank of stromboli volcano during and after the 2002–2003 eruption," *Bull. Volcanol.*, vol. 70, no. 6, pp. 703–715, Apr. 2008, doi: 10.1007/s00445-007-0162-1.

[2] J. Chandler, "Effective application of automated digital photogrammetry for geomorphological research," *Earth Surf. Process. Landforms*, vol. 24, no. 1, pp. 51–63, Jan. 1999, doi: 10.1002/(SICI)1096-9837(199901)24:1<51::AID-ESP948>3.0.CO.2-H.

[3] O. Tziavou, S. Pytharouli, and J. Souter, "Unmanned aerial vehicle (UAV) based mapping in engineering geological surveys: Considerations for optimum results," *Eng. Geol.*, vol. 232, pp. 12–21, Jan. 2018, doi: 10.1016/j.enggeo.2017.11.004.

[4] M. J. Westoby, J. Brasington, N. F. Glasser, M. J. Hambrey, and J. M. Reynolds, "'Structure-from-motion' photogrammetry: A low-cost, effective tool for geoscience applications," *Geomorphology*, vol. 179, pp. 300–314, Dec. 2012, doi: 10.1016/j.geomorph.2012.08.021.

[5] R. A. Haugerud, D. J. Harding, S. Y. Johnson, J. L. Harless, C. S. Weaver, and B. L. Sherrod, "High-resolution lidar topography of the Puget Lowland, Washington—A bonanza for Earth science," *GSA Today*, vol. 13, no. 6, pp. 4–10, 2003, doi: 10.1130/1052-5173(2003)13<0004:HLTOTP>2.0.CO.2.

[6] K. L. Cook, "An evaluation of the effectiveness of low-cost UAVs and structure from motion for geomorphic change detection," *Geomorphology*, vol. 278, pp. 195–208, Feb. 2017, doi: 10.1016/j.geomorph.2016.11.009.

[7] S. W. Son, D. W. Kim, W. G. Sung, and J. J. Yu, "Integrating UAV and TLS approaches for environmental management: A case study of a waste stockpile area," *Remote Sens.*, vol. 12, no. 10, p. 1615, May 2020, doi: 10.3390/rs12101615.

[8] C. Zhang and J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: A review," *Precis. Agricult.*, vol. 13, no. 6, pp. 693–712, Dec. 2012, doi: 10.1007/s11119-012-9274-5.

[9] F. J. Aguilar, F. Agüera, M. A. Aguilar, and F. Carvajal, "Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy," *Photogramm. Eng. Remote Sens.*, vol. 71, no. 7, pp. 805–816, Jul. 2005, doi: 10.14358/PERS.71.7.805.

[10] M. R. James, S. Robson, S. d'Oleire-Oltmanns, and U. Niethammer, "Optimising UAV topographic surveys processed with structure-from-motion: Ground control quality, quantity and bundle adjustment," *Geomorphology*, vol. 280, pp. 51–66, Mar. 2017, doi: 10.1016/j.geomorph.2016.11.021.

[11] M. Wulder, "Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters," *Prog. Phys. Geography, Earth Environ.*, vol. 22, no. 4, pp. 449–476, Dec. 1998, doi: 10.1177/030913339802200402.

[12] M. R. James and S. Robson, "Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application," *J. Geophys. Res., Earth Surf.*, vol. 117, no. 3, pp. 1–17, Sep. 2012, doi: 10.1029/2011JF002289.

[13] S. Manfreda *et al.*, "On the use of unmanned aerial systems for environmental monitoring," *Remote Sens.*, vol. 10, no. 4, p. 641, Apr. 2018, doi: 10.3390/rs10040641.

[14] S. R. Herwitz *et al.*, "Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support," *Comput. Electron. Agricult.*, vol. 44, no. 1, pp. 49–61, Jul. 2004, doi: 10.1016/j.compag.2004.02.006.

[15] J. N. Callow, S. M. May, and M. Leopold, "Drone photogrammetry and KMeans point cloud filtering to create high resolution topographic and inundation models of coastal sediment archives," *Earth Surf. Processes Landforms*, vol. 43, no. 12, pp. 2603–2615, Sep. 2018, doi: 10.1002/esp.4419.

[16] I. Colomina and P. Molina, "Unmanned aerial systems for photogrammetry and remote sensing: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 92, pp. 79–97, Jun. 2014, doi: 10.1016/j.isprsjprs.2014.02.013.

[17] K. K. Singh and A. E. Frazier, "A meta-analysis and review of unmanned aircraft system (UAS) imagery for terrestrial applications," *Int. J. Remote Sens.*, vol. 39, nos. 15–16, pp. 5078–5098, Aug. 2018, doi: 10.1080/01431161.2017.1420941.

[18] A. Eltner and D. Schneider, "Analysis of different methods for 3D reconstruction of natural surfaces from parallel-axes UAV images," *Photogramm. Rec.*, vol. 30, no. 151, pp. 279–299, Sep. 2015, doi: 10.1111/phor.12115.

[19] M. R. James and S. Robson, "Mitigating systematic error in topographic models derived from UAV and ground-based image networks," *Earth Surf. Processes Landforms*, vol. 39, no. 10, pp. 1413–1420, Aug. 2014, doi: 10.1002/esp.3609.

[20] A. Lucieer, S. M. D. Jong, and D. Turner, "Mapping landslide displacements using structure from motion (SfM) and image correlation of multi-temporal UAV photography," *Prog. Phys. Geography, Earth Environ.*, vol. 38, no. 1, pp. 97–116, Feb. 2014, doi: 10.1177/0309133313515293.

[21] G. Tmušić *et al.*, "Current practices in UAS-based environmental monitoring," *Remote Sens.*, vol. 12, no. 6, p. 1001, Mar. 2020, doi: 10.3390/rs12061001.

[22] R. H. Bewley, S. P. Crutchley, and C. A. Shell, "New light on an ancient landscape: Lidar survey in the stonehenge world heritage site," *Antiquity*, vol. 79, no. 305, pp. 636–647, Sep. 2005.

[23] P. Novakovi, M. Hornak, and M. Zachar, "3D digital recording of archaeological," *Architectural Artistic Heritage*, pp. 33–51, May 2017.

[24] M. Reich, M. Wiggenhagen, and D. Muhle, "Filling the holes-potential of UAV-based photogrammetric façade modelling," *Tagungsband Des*, vol. 15, pp. 147–156, May 2012.

[25] A. Koutsoudis, B. Vidmar, G. Ioannakis, F. Arnaoutoglou, G. Pavlidis, and C. Chamzas, "Multi-image 3D reconstruction data evaluation," *J. Cultural Heritage*, vol. 15, no. 1, pp. 73–79, Jan. 2014, doi: 10.1016/j.culher.2012.12.003.

[26] A. Koutsoudis, G. Ioannakis, B. Vidmar, F. Arnaoutoglou, and C. Chamzas, "Using noise function-based patterns to enhance photogrammetric 3D reconstruction performance of featureless surfaces," *J. Cultural Heritage*, vol. 16, no. 5, pp. 664–670, Sep. 2015, doi: 10.1016/j.culher.2015.01.008.

[27] A. Jalandoni, I. Domingo, and P. S. C. Taçon, "Testing the value of low-cost structure-from-motion (SfM) photogrammetry for metric and visual analysis of rock art," *J. Archaeol. Sci., Rep.*, vol. 17, pp. 605–616, Feb. 2018, doi: 10.1016/j.jasrep.2017.12.020.

[28] F. Remondino, M. G. Spera, E. Nocerino, F. Menna, and F. Nex, "State of the art in high density image matching," *Photogramm. Rec.*, vol. 29, no. 146, pp. 144–166, Jun. 2014, doi: 10.1111/phor.12063.

[29] L. Roth, A. Hund, and H. Aasen, "PhenoFly planning tool: Flight planning for high-resolution optical remote sensing with unmanned areal systems," *Plant Methods*, vol. 14, no. 1, p. 116, Dec. 2018, doi: 10.1186/s13007-018-0376-6.

[30] S. Berquist *et al.*, "A new aerial photogrammetric survey method for recording inaccessible rock art," *Digit. Appl. Archaeology Cultural Heritage*, vol. 8, pp. 46–56, Mar. 2018, doi: 10.1016/j.daach.2018.03.001.

[31] J. Barlow, J. Gilham, and I. Ibarra Cofré, "Kinematic analysis of sea cliff stability using UAV photogrammetry," *Int. J. Remote Sens.*, vol. 38, nos. 8–10, pp. 2464–2479, May 2017, doi: 10.1080/01431161.2016.1275061.

[32] M. Jaud *et al.*, "UAV survey of a coastal cliff face–selection of the best imaging angle," *Measurement*, vol. 139, pp. 10–20, Jun. 2019, doi: 10.1016/j.measurement.2019.02.024.

[33] R. Wackrow and J. H. Chandler, "A convergent image configuration for DEM extraction that minimises the systematic effects caused by an inaccurate lens model," *Photogramm. Rec.*, vol. 23, no. 121, pp. 6–18, Mar. 2008, doi: 10.1111/j.1477-9730.2008.00467.x.

[34] P. Martínez-Carricondo, F. Agüera-Vega, and F. Carvajal-Ramírez, "Use of UAV-photogrammetry for quasi-vertical wall surveying," *Remote Sens.*, vol. 12, no. 14, Art. no. 14, Jan. 2020, doi: 10.3390/rs12142221.

[35] USGS National UAS Project Office. (Mar. 2017). *USGS Agisoft PhotoScan Workflow*. Accessed: Oct. 23, 2019. [Online]. Available: https://uas.usgs.gov/nupo/pdf/USGSAgisoftPhotoScanWorkflow.pdf

[36] A. Lucieer. (2011). *Algorithms used in Photoscan*. Accessed: Apr. 29, 2020. [Online]. Available: https://www.agisoft.com/forum/index.php?topic=89.0

[37] CloudCompare. (2016). *CloudCompare*. Accessed: Mar. 11, 2020. [Online]. Available: http://www.cloudcompare.org/

[38] D. D. Lichti, S. J. Gordon, and T. Tipdecho, "Error models and propagation in directly georeferenced terrestrial laser scanner networks," *J. Surv. Eng.*, vol. 131, no. 4, pp. 135–142, 2005, doi: 10.1061/(ASCE)0733-9453(2005)131:4(135).

[39] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *Int. J. Comput. Vis.*, vol. 13, no. 2, pp. 119–152, Oct. 1994, doi: 10.1007/BF01427149.

[40] QGIS. (2020). *QGIS Geographic Information System Open Source Geospatial Foundation Project*. [Online]. Available: https://qgis.org/

[41] G. Forlani, E. Dall'Asta, F. Diotri, U. M. D. Cella, R. Roncella, and M. Santise, "Quality assessment of DSMs produced from UAV flights georeferenced with on-board RTK positioning," *Remote Sens.*, vol. 10, no. 2, p. 311, Feb. 2018, doi: 10.3390/rs10020311.

[42] K. Kraus, *Photogrammetry: Geometry from Images and Laser Scans*. Berlin, Germany: Walter de Gruyter, 2011.

[43] Agisoft LLC. (2019). *Agisoft Metashape User Manual: Professional Edition, Version 1.5*. Accessed: Jan. 16, 2019. [Online]. Available: https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf

[44] A. Pasumansky. (2018). *Topic: Masking the Sky Using Background Photos*. Accessed: Apr. 23, 2020. [Online]. Available: https://www.agisoft.com/forum/index.php?topic=9231.msg43077#msg43077

[45] J. Dandois, M. Olano, and E. Ellis, "Optimal altitude, overlap, and weather conditions for computer vision UAV estimates of forest structure," *Remote Sens.*, vol. 7, no. 10, pp. 13895–13920, Oct. 2015, doi: 10.3390/rs71013895.

[46] E. Seifert *et al.*, "Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images," *Remote Sens.*, vol. 11, no. 10, p. 1252, May 2019, doi: 10.3390/rs11101252.

[47] C.-S. Lin, S.-H. Chen, C.-M. Chang, and T.-W. Shen, "Crack detection on a retaining wall with an innovative, ensemble learning method in a dynamic imaging system," *Sensors*, vol. 19, no. 21, p. 4784, Nov. 2019, doi: 10.3390/s19214784.

[48] S. Daftry, C. Hoppe, and H. Bischof, "Building with drones: Accurate 3D facade reconstruction using MAVs," in *Proc. IEEE Int. Conf. Robot. Automat.*, Seattle, WA, USA, May 2015, pp. 3487–3494, doi: 10.1109/ICRA.2015.7139681.

[49] A. Pasumansky. (2017). *Topic: About 'Build Orthomosaic*. Accessed: May 20, 2017. [Online]. Available: https://www.agisoft.com/forum/index.php?topic=7000.0

[50] M. J. Henriques, A. Fonseca, D. Roque, J. N. Lima, and J. Marnoto, "Assessing the quality of an UAV-based orthomosaic and surface model of a breakwater," presented at the FIG Congr., 2014. [Online]. Available: https://mycoordinates.org/assessing-the-quality-of-an-uav-based-orthomosaic-and-surface-model-of-a-breakwater/

[51] Y.-H. Tu, S. Phinn, K. Johansen, A. Robson, and D. Wu, "Optimising drone flight planning for measuring horticultural tree crop structure," *ISPRS J. Photogramm. Remote Sens.*, vol. 160, pp. 83–96, Feb. 2020, doi: 10.1016/j.isprsjprs.2019.12.006.

Yu-Hsuan Tu (Member, IEEE) received the Ph.D. degree in Earth and environmental sciences from the University of Queensland, Brisbane, Australia, in 2019.

Since 2019, he has been a Post-Doctoral Fellow with the King Abdullah University of Science and Technology (KAUST) Thuwal, Saudi Arabia. His research focuses on high-resolution remote sensing, UAVs, vegetation mapping, and spatial analysis.

Kasper Johansen received the Ph.D. degree in remote sensing from the University of Queensland, Brisbane, Australia, in 2007.

Since 2017, he has been a Research Scientist with KAUST, Thuwal, Saudi Arabia, where he is an Associate Editor of Frontiers in Remote Sensing. His main research interests are focused on processing and analyzing high spatial resolution remotely sensed multispectral, hyper-spectral, LiDAR and thermal image data acquired from satellite, airborne and UAV-based sensors to extract ecologically meaningful information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TU *et al.*: COMBINING NADIR, OBLIQUE, AND FAÇADE IMAGERY ENHANCES RECONSTRUCTION OF ROCK FORMATIONS
13

Bruno Aragon received the M.S. degree in environmental engineering from KAUST, Thuwal, Saudi Arabia, in 2015, where he is pursuing the Ph.D. degree.

His research focus is in the field of precision agriculture and evaporation estimation using remotely sensed data, particularly with Cubesats.

Samir K. M. Al-Mashharawi received the M.S. degree in environmental science and engineering from KAUST, Thuwal, Saudi Arabia, in 2011.

Since 2011, he has been a Research Engineer with KAUST. His research is focused on the novel use of field-based technologies for enhanced agricultural monitoring.

Bonny M. Stutsel received the Ph.D. degree in geography from The University of Western Australia, Perth, Australia in 2019.

Since 2019, she has been a Post-Doctoral Fellow with KAUST, Thuwal, Saudi Arabia. Her research explores how in-field, proximal, and remote sensing platforms can be applied to precision agriculture to observe plant stress, facilitate phenotyping, and reduce inputs. She has a broad interest in how multidisciplinary research can translate into smarter farming practices and inform water-food nexus policy.


Jiale Jiang received the Ph.D. degree in surveying and mapping from the China University of Geosciences, Beijing, China, in 2017.

Since 2019, she has been a Post-Doctoral Fellow at KAUST, Thuwal, Saudi Arabia. Her research is focused on spatiotemporal image fusion and crop growth monitoring.

Yoseline Ángel received the M.S. degree in geomatics from the National University of Colombia, Bogotá, Colombia, in 2012. She is pursuing the Ph.D. degree in environmental science and engineering with KAUST, Thuwal, Saudi Arabia.

Her research is focused on hyperspectral remote sensing of biophysical metrics of crops.
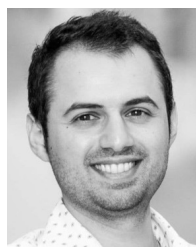
Matteo G. Ziliani received the M.S. degree in civil and environmental engineering from the Polytechnic University of Milan, Milan, Italy, in 2015. He is pursuing the Ph.D. degree in environmental science and engineering with KAUST, Thuwal, Saudi Arabia.

His research is focused on crop yield forecasting methods integrated with UAV and satellite-based remote sensing applications.

Omar A. López Camargo received the B.S. degree in electronic engineering from the National University of Colombia, Bogotá, Colombia, in 2020. He is pursuing the M.S./Ph.D. degree in environmental sciences and engineering with KAUST, Thuwal, Saudi Arabia.

His research is focused on satellite systems and robotics.

Matthew F. McCabe received the Ph.D. degree in civil and environmental engineering from the University of Newcastle, Callaghan, Australia, in 2003.

He is a Professor of Remote Sensing and Water Security with KAUST, Thuwal, Saudi Arabia. His research explores a range of multidisciplinary issues around water and food security, climate change impacts, precision agriculture, and water resources monitoring and modeling. He has been involved in a number of international climate research and coordination activities (G20, WCRP, GCOS, and GEWEX), and is an inaugural Specialty Chief Editor for Frontiers in Artificial Intelligence.