A Proposal for the Standardized Reporting of Error and Paradata Regarding Structure from Motion (SfM) 3D Models Used in Recording and Consolidating Archaeological Architecture

Mario Borrero and Luke R. Stroth

ABSTRACT

In the past decade, archaeologists have increasingly made use of photogrammetry, the process of creating 3D models from photographs, in a variety of field and lab settings. We argue that we must, as a discipline, develop a consistent methodology to ensure that 3D models are held to a consistent standard, including not only photographic protocol but also the documentation of model accuracy using an agreed-upon measure. To help develop this discussion, we present our system for incorporating photogrammetry into the documentation of architecture. This technique was developed at the site of Nim Li Punit, Belize, in 2018. Excavating architecture involves documenting the pre-excavated building, liberating overburden, documenting all in situ construction (including wall fall, fill stones, and standing architecture), drawing consolidated architecture, and documenting the final state of the post-excavated buildings. The generation of 3D models greatly assisted in all facets of the excavation, documentation, analysis, and consolidation processes. To ensure that our models were accurate, we documented the reprojection error and final model horizontal distortion to assess the quality of the model. We suggest that documenting both forms of error should become standard practice in any discussion of archaeological applications of photogrammetry.

Keywords: photogrammetry, architecture, 3D modeling, Structure from Motion (SfM), Maya archaeology, field methods

En la última década, los arqueólogos han recurrido cada vez más al uso de la fotogrametría, el proceso en el cual crean modelos tridimensionales (3D) con fotografías. Este proceso se da una variedad de lugares como en el campo y/o laboratorio. Discutimos que, como disciplina, debemos desarrollar una metodología consistente para garantizar que los modelos 3D se mantengan en un estándar consistente, que incluya no solo el protocolo fotográfico y que la documentación sea precisa y utilice las medidas acrodadas en la metodología aprobada. Para ayudar a desarrollar esta discusión, presentamos nuestro sistema para incorporar fotogrametría en la documentación de la arquitectura. Esta técnica se desarrolló en el sitio arqueológico de Nim Li Punit, Belice, en 2018. La excavación de arquitectura consiste en documentar el edificio preexcavado, liberar la sobrecarga, documentar toda la construcción in situ (incluida la caída de muros, piedras de relleno y arquitectura en pie), dibujos de la arquitectura consolidada, y documentación del estado final de los edificios post-excavación. La generación de modelos 3D ha ayudado mucho en todas las facetas de los procesos de excavación, documentación, análisis y consolidación. Para garantizar que nuestros modelos fueran precisos, documentamos el error de reproyección para evaluar la calidad de la alineación y los cambios en las distancias entre los puntos de control ya conocidos medidos en el mundo real, y las medidas realizadas en el mismo punto en el espacio digital. Nuestra meta será la de evaluar el grado de distorsión en el proyecto final. Sugerimos que documentar ambas formas de error se convierta en una práctica estándar en futuras discusiones sobre aplicaciones arqueológicas utilizando la fotogrametría.

Palabras clave: fotogrametría, arquitectura, modelado en 3D, estructura a partir del movimiento (Structure from Motion; SfM), arqueología maya, técnica de excavación

Archaeologists have digitally recorded artifact provenance since the 1970s (Forte et al. 2012; McWilliams et al. 2005; Pavlidis 2006; Reilly 1991). These methods are now the standard for measurement and recording. Structure from Motion photogrammetry (SfM), the creation of 3D models from overlapping photographs, is one such method that has become increasingly popular (Douglass et al. 2020).
After over a decade of SFM in the field, researchers have expressed concern about best practices of reporting and assessing the model accuracy (Antón et al. 2018; Barbasiwicz et al. 2018; Fei Peng et al. 2017; Richards-Rissetto 2017; Sapirstein 2016; SAPirstein and Murray 2017). In this article, we propose a set of best practices to ensure that SFM practitioners can evaluate the accuracy of published models. These methods should assist the researcher in rapidly evaluating the trustworthiness of a model as an accurate representation of real-world archaeological sites, features, and artifacts. We anchor our discussion of these practices in how we use SFM to facilitate the consolidation of collapsed architecture at the Classic Maya site of Nim Li Punit, Toledo District, Belize. We present our workflow, describe the advantages of the technique, and specify how to assess the accuracy of the 3D models and all subsequent representations of these ancient buildings. The core of our suggested best practices includes reporting (Table 1; Supplemental Table 1) the reprojection error, horizontal distortion, and all paradata associated with model generation.

Traditional hand-drawn methods (Figure 1) remain a staple of architectural illustration, particularly in Mesoamerican archaeology. Horizontal measurements are taken from the x- and y-axes of the unit with tape measure and line level. Z measurements are made with a plumb bob, producing a scaled drawing displaying features of interest. Although it is the most cost-effective option, it is time consuming, and neither the rate of error nor accuracy can be reported or independently evaluated (Borrero et al. 2019). Hand-drawn illustrations are interpretations; some elements may be cut or distorted, along with the inherent distortion of recreating a 3D object as a 2D surface (Douglas et al. 2015; Lemmi et al. 2010). Accurate spatial data is critical to any analysis. Our experience with SFM has provided an alternative recording technique in which the accuracy of the final product is known.

3D recording is at the forefront of data acquisition, storage, and visualization. These models are superior for presenting volumetric information, visualizing surrounding landscapes, and assessing construction materials and other ephemeral features (Fritz et al. 2016; Koutsoudis et al. 2014; Plets et al. 2012). 3D data are necessary to document, analyze, and reconstruct ancient structures. In a recent article, Sapirstein and Murray (2017) called for establishing best practices for photogrammetry in archaeology. We agree with their perception of an “urgent need” (Sapirstein and Murray 2017:348) for an academic standard. Similar to an industry standard, we hope to contribute to this developing discourse regarding a series of specific requirements for SFM within archaeological practice. Like the authors, we advocate for reporting standardized errors and metadata. Furthermore, we advocate for reporting the reprojection error and final horizontal distortion (the difference between real-world measurements of ground control points and those in the model) and including the paradata regarding decisions made in the construction of the model.

**SFM PHOTOGRAHAMETRY AT NIM LI PUNIT**

Nim Li Punit (AD 150–830/850), Toledo District, Belize (Figure 2), was a small polity that exhibited political connections to powerful centers such as Caracol, Altun Ha, and Quiriguá or Copán in its hieroglyphic record and imported artifacts (Braswell 2017; Helmke et al. 2018; Prager and Braswell 2016). Limited test excavations and government consolidation work were carried out throughout the 1980s and ’90s (Hammond et al. 1999). The Toledo Regional Interaction Project (TRIP), directed by Geoffrey Braswell (University of California, San Diego), has worked at Nim Li Punit since 2010 (Braswell 2017). We first implemented SFM photogrammetry during the 2018 field season (Borrero et al. 2019).

In June 2018, we presented on the way we use SFM photogrammetry (Borrero et al. 2019). Many other members of the Belizean archaeological community had likewise used SFM photogrammetry to reproduce individual artifacts (e.g., Shurik 2019; Skaggs et al. 2018) or to assist in documenting and illustrating stratigraphic profiles (Grauer et al. 2019). These are worthy pursuits, facilitating the analysis of material culture while the original artifacts remain in the country of origin. Our contribution to this existing discourse is to emphasize standardized reporting of error as part of best practices and to stress the importance of 3D modeling to accurate architectural reconstruction.

Architectural illustrations are a staple of Maya archaeology (Halperin and Garrido 2020; Halperin et al. 2019; Harrison-Buck 2012; Houston 1998). They are a standard way of presenting the results of excavations to the academic community and the public. Accurate representation of architecture allows us to identify contemporary stages of construction across multiple buildings, distinguish between regional traditions, and assess cultural logic in building patterns (Ashmore and Sabloff 2002; Hohmann-Vogrin 2006; Lominy 2006; von Schwerin 2011; Webster 1998). Photographs might also be used for this purpose (e.g., Powis et al. 2019), but illustrations filter out unnecessary detail. The architectural history at Nim Li Punit is complex, with multiple overlapping construction phases (Braswell 2017). There is evidence of deconstruction and looting toward the end of the occupation of the site (Braswell et al. 2019). Accurately documenting these architectural phenomena is critical to uncovering the cultural history of Nim Li Punit and the southern Belize region, but there is no way to gauge error with hand-drawn, tape and plumb bob depictions. Sapirstein and Murray ask that “digital products be held to the same standards of precision, accuracy, and sustainability as the analog recording techniques they replace” (2017:349). In this article, we

---

**Table 1. Total Number of Projects and Photographs, Average Number of Photos per Feature and Unit, Average RMS Error, and Total and Average File Size.**

<table>
<thead>
<tr>
<th>Number Projects (features)</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Projects (units)</td>
<td>20</td>
</tr>
<tr>
<td>Total Number Photos (features)</td>
<td>290</td>
</tr>
<tr>
<td>Total Number Photos (units)</td>
<td>987</td>
</tr>
<tr>
<td>Average Photos per Feature</td>
<td>32.22</td>
</tr>
<tr>
<td>Average Photos per Unit*</td>
<td>22.43</td>
</tr>
<tr>
<td>Average RMS Error (all)</td>
<td>0.1193</td>
</tr>
<tr>
<td>Total File Size (GB)</td>
<td>150.326</td>
</tr>
<tr>
<td>Average Size of Model (GB)</td>
<td>4.698</td>
</tr>
</tbody>
</table>

*Seven sessions documented multiple 2 × 2 m units. The average number of photos per unit is divided by the total number of 2 × 2 m units that were documented (n = 44) rather than the number of projects documenting units.
argue for a set of best practices that go beyond what hand-mapping can offer: quantifiable error.

SfM photogrammetry employs multiview stereo algorithms that estimate the location of the camera to model the 3D surface geometry of the subject. Using matching points from overlapping images taken at multiple angles (Figure 3a), the software produces different lines of sight between the camera and the subject. Triangulating between them generates points on the surface of the subject in virtual space. SfM programs can produce highly accurate and precise 3D models of archaeological material. Close-range photogrammetry refers to image-capturing within 1–2 m of a subject (Lerma et al. 2010). This technique is commonly used to document historical monuments because it captures their structural details and textures (Koeva 2004; Yilmaz et al. 2007). Accurate measurements can be taken from these models, which facilitates independent review of the data. In the field, SfM photogrammetric measurements are a faster and easier alternative to traditional geodetic measurements, and they do not disrupt excavation (Barbasiewicz et al. 2018; Koutsoudis et al. 2014; Myers and Badillo 2019). These models can be converted into precise 3D maps, orthophotos (planar views of a 3D scene that maintain accurate scale along the x- and y-axes of the image), and traditional scaled 2D plan drawings. We employed SfM photography during our 2018 field season at Nim Li Punit, Belize, to efficiently and accurately record the architectural histories of the structures we excavated (Borrero et al. 2019; Braswell et al. 2019).

Previously, graph-paper drawings were scanned and traced in Adobe Illustrator, introducing various stages of potential human error. We were able to minimize the error of in-field measurements through the direct import of scaled orthophotos to Illustrator. An important point is that the accuracy of measurements within any model is derived from the way the physical measurements were taken (Fei Peng et al. 2017)—our ground control points (GCPs; coded targets, detectable by Metashape, that were printed and laminated, sensu Sapirstein [2016:5]) were set up using a tape measure and a meter scale bar. The corners of each unit were tied to GCPs tied to the UTM coordinate system. We chose to place individual coded targets for our GCPs because our subjects were complex and irregular. This required flexibility to best capture the shape of the fallen architecture. Digitizing individual artifacts or a smaller excavated context requires finer control of distance between points, such as coded targets fixed to a scale bar at a constant distance. Given the scale of our digitized subject, this level of accuracy was not necessary. Ultimately, horizontal distortion is determined by the number of and distance between GCPs. This should be addressed by the researchers prior to excavation in order to keep accuracy consistent between models (Fei Peng et al. 2017).

Measuring and Reporting Error

Randles and colleagues (2010) and similar studies from the automotive industry demonstrate that photogrammetric
measurements taken from 3D models of vehicles fall within acceptable error ranges for their industry’s accident reconstruction. They proposed the use of coded targets set at known distances to assist with the modeling process and increase precision. Documentation of error is important for understating the quality of a model. From an industrial engineering perspective, this speaks to the “fitness of use[,]” or how appropriate the model is for a given analysis and how precisely it represents the real-world subject (Chrisman 1983). Several industries, including automotive manufacturing, crash and accident assessment, aerospace, wind energy systems, engineering and construction, and cartographers and surveyors, have adopted photogrammetry and fully accept digital 3D as a measurement tool. Many of these professionals go to great lengths to reduce the error in their measurements, emphasizing the reduction of deviation between positional measurements and the “ground truth” (Niederheiser et al. 2016:165)

Reporting error is not a sign of weakness in other disciplines; error estimates provide crucial information that must be preserved to
properly interpret a model. Error is a fundamental dimension of data that must be acknowledged. Therefore, we promote the frank discussion of model accuracy regarding the dissemination of 3D models for archaeological research.

Our acceptable level of accuracy for capturing a single $2 \times 2$ m unit at less than 5 cm horizontal distortion in any direction (2.5% error). This scaled according to the size of the project (Sapirstein and Murray 2017); when documenting a $10 \times 4$ m structure, the acceptable level of error was less than 50 cm. We consistently used reprojection error of 0.15 (as reported in Metashape) before the model was processed. Reprojection error (RE) is the difference between a projected point and a measured one. High RE may stem from issues with initial ground control point placement, scaling within software, poor camera geometry during the capturing session, lens distortion, and, especially, alignment within the camera body. This can be overcome through experience, control of camera settings, maintaining strong lines between the camera and subject, and considering coverage and angle of view during photography. The lower this value, the more trustworthy the estimated points of the object’s geometry of representing the real-world object.

FIGURE 3. Processing photographs to create 3D models: (a) dense point cloud of Operation 5 Unit 48U Lot 1, illustrating the location of the camera in relation to the unit; (b) editing in the RAW photographs in Adobe Lightroom for color correction and white balancing.
WORKFLOW FOR SFM DOCUMENTATION AND RECONSOLIDATION OF ARCHITECTURE

The SFM photogrammetry workflow must incorporate research design from the beginning, writing a data management plan to keep methods consistent between projects and ensure that there are long-term data storage solutions (Fernandez 2019; Myers and Badillo 2019). The degree to which similar research designs are followed will depend on the purpose for 3D modeling for the proposed project. If the intention is to record accurately and preserve in a measurable manner the archaeological site, building, or artifact, then rigorous measuring protocols must be implemented. Data management and workflow are contingent on the project parameters, the hypothesis being tested, and the expected file size. The requirements for a landscape survey will be quite different from documenting individual artifacts (Fei Peng et al. 2017; Sapirstein and Murray 2017). The procedure we present was implemented in its current form in our 2019 field season. We used the program Metashape (formerly PhotoScan), produced by Agisoft. We chose it because the lead author was trained in the program and it is well represented in the literature. Metashape has been shown to have a minimal difference in measurement accuracy as compared to other programs such as Pix4D (Barbasiewicz et al. 2018; Burns and Delpate 2017).

TRIP SfM Workflow

- Acquire photographs at the scene
  - Document metadata (unit/feature, date, photographer, measured distance between GCPs, presence/absence of color checker/white balance card)
  - Document paradata (camera used, camera settings [e.g., ISO, shutter speed, F-stop, etc.], method of lighting [natural or artificial, reflectors, presence of cover over the subject, etc.], file format in which photographs were saved)
- Digitally process photographs
  - Document any preprocessing editing of images (e.g., color correction, editing saturation/contrasts, removal of blurry images, etc.)
- Generate 3D model
  - Document initial RE and final RE after the model was edited, the way the model was scaled, and the editing of points.
- Texturing and Visualization
  - Export models and present them in post-processing programs. Produce a scaled illustration using an orthophoto.
- Reconsolidate
  - Port the model to a mobile phone/tablet for field display. Use the accurate model to guide consolidation that showcases the history of the building.

Materials for SfM in the Field

Our field kit included a Canon EOS 80D DSLR with an 18–55 mm EFS lens, two 64 GB digital SD cards, an X-Rite Passport Color Checker and white balance card, 12 mm coded targets printed on white stock and laminated, a telescopic monopod for overhead shots, a white board to record metadata, a 1 m scale bar, and a north arrow. We locked manual focus and f/8 stop across the photo-taking session to increase model accuracy and reduce processing time (Sapirstein and Murray 2017:341). The ISO and exposure were set according to lighting conditions and fixed for the duration of the photography session. To process the photographs, we used a 2015 MacBook Pro with a 3.1 GHz Intel processor and 16 GB of RAM, running Metashape’s Professional Edition (version 1.5.0.7492), and a 2017 Dell Inspiron 15 7000 Series Gaming 2.5 GHz Intel processor and 8 GB of RAM, running Metashape’s Standard Edition. Both versions can generate models, but only the Professional Edition can produce orthophotos and scale the model. Data storage and transfer were accomplished with two external hard drives.

Data Acquisition

We first defined the study space and recorded metadata, including the unit(s)/feature(s), distance between GCPs, and date. Each session constituted a single “project.” For architectural illustration, we created a new project for each unit or series of units when the overburden was removed, exposing the architecture. Each project included a scale bar, a north arrow, the metadata written on a whiteboard, and a color-checker/white balance. We always included the four corners of the exterior polygon of the project, typically a single 2 × 2 m unit or series of adjacent units, and we placed GCPs at predetermined distances. The corners of each unit are tied into the grid established by Braswell in 2012 that is georeferenced to the UTM coordinate system: Zone 16 North, WGS 1984 datum (Braswell and Daniels 2013). The subjects were photographed along a central axis, an angled right and left axis relative to the first, above the head, at eye level, and below the waist (Figure 3a). Consistency in camera angle and reduction of distance between GCPs will minimize distortion (Fei Peng et al. 2017; Ishida 2017). Typical 2 × 2 m units required 22–23 photographs (Table 1). Particularly intricate lots with a high volume of artifacts required additional photography and close-up shots. Planar overhead shots, with all four corners of the unit in frame, were taken with a monopod. Several photos were taken overlapping with other units to provide tie-in points across the excavation area between the different subject units.

Good photographs should be properly exposed, be evenly illuminated, be focused, contain good contrast and texture, and have minimal image compression. Best results come by keeping camera settings (i.e., ISO, aperture, and shutter speed) consistent (Ishida 2017; Pavlidis et al. 2006; Porter et al. 2016). All photos were taken using manual focus and a consistent focal length. Our photographs were taken in RAW format. This file format reduces the amount of on-board processing by the proprietary systems of the camera. We stored our images as both JPEG and CR2 files (the proprietary RAW format for Cannon). By processing RAW images through such camera file conversion software as Adobe Lightroom, a variety of image issues—including over- or underexposure—can be corrected. This produces better results than using the JPEGs edited following factory defaults. Unfortunately, the file size of RAW images is considerably larger than other formats.

We recorded the paradata—the choices made throughout the project that introduce bias and affect the outcome. The final accuracy of the model can be quantitatively assessed using reprojection error and final horizontal distortion, but decisions will inform the final interpretation of the model, such as those made...
during the data collection process (e.g., photographic variables, camera settings, choice of scene) and edits made during the modeling process (e.g., decimation of mesh for online display, removal of points, post-processing additions made to a model). This allows others to evaluate model accuracy as it represents a real subject (Richards-Rissetto and von Schwerin 2017). A balance must be struck between transparency and ease of reporting. These decisions may not be relevant if the only goal is to create a 3D model for museum or online display. In contrast, any model that is used for quantitative analysis must be subject to independent review. Including the decisions made during model processing in the supplemental material would satisfy this need.

Data Processing

Having used the color-checker/white balance, we were able to process our RAW images in Adobe Lightroom to have uniform lighting without changing the pixel value within the photograph (Figure 3b). This was often necessary due to the variable shade beneath the tree canopies. Following bulk processing, redundant, out-of-focus, or unnecessary pictures were removed. This is the first piece of paradata to be reported. The corrected photographs were imported into Metashape. We oriented (called “alignment” in Metashape) photos with generic preselection turned off. Generic Preselection first finds pairs of overlapping photographs by attempting to align them at lower resolution settings. Turning off Generic Preselection takes longer but provides a higher resolution alignment (Agssoft 2019). If the reprojection error exceeded our chosen maximum of 0.15, it was re-run with problematic (out of focus, in shadow, or unaligned) pictures removed. Reprojection error is a geometric error corresponding to the image distance between a measured point and where the point was projected to fall. This ratio is used to quantify how closely an estimated 3D point re-creates a point’s true projection—measuring the mismatch between the location of the generated point and the original subject. Error during alignment can stem from poor camera geometry, lack of GCPs/scale, and lens distortion.

Aligning the photographs creates a sparse point cloud. Problematic points and unnecessary background detail, such as trees, buckets, and the photographer’s feet, can be edited out at this stage. This qualitative selection of irrelevant features again introduces bias. Another option for the Metashape user is the gradual selection tool, where all points that fall above or below a certain threshold for RE can be selected and removed. After the points have been removed, the cameras can be reoriented to further reduce the RE.

Root mean square (RMS) error is an assessment of the accuracy of the model, typically conducted through a standard geocorrection error measurement. The acceptable level of error depends on the project. RMS error values are defined in terms of input image pixel size. A good architectural photogrammetry project should have an RMS residual value of less than 3, with a value of less than 1.5 being excellent (McWilliams et al. 2005). Individual points that are not attached to the model, or with a high RMS or RE, can be removed to improve overall model accuracy. Overall, accuracy relies heavily on an even distribution of control points across the captured scene, and it is generally good practice to include many (Fei Peng et al. 2017; Porter et al. 2016). We believe these errors should be reported as an academic standard (Supplemental Figure 1). The 3D model is an interpretation of the subject, the product of a series of choices made by the analyst (Richards-Rissetto and von Schwerin 2017). Ethical modeling should document these user decisions that comprise the paradata (Fernandez 2019; Turco et al. 2019). Both quantitative measures of error (reprojection error for data going in; horizontal distortion for error coming out) and subjective decisions (cropping the image because it looks nicer) should be reported.

3D Model Generation

In the field, Image Matching (“Alignment”), Dense Point Cloud, and Mesh were all done on Medium (Figures 4a and 4b) to balance processing time, file size, and accuracy. Most modeling programs generate a wireframe (“mesh”) by triangulating between different points of the dense point cloud. 3D wireframe models do provide the true shape of the building but appear akin to a line drawing (Figures 4c and 4d). To arrive at a more realistic impression of the structure, it is necessary to drape texture over the wireframe (Figures 4e and 4f). Texture is applied from selected images generated in the capturing phase. These can be chosen manually or automatically based on the images with best resolution.

Although we did not produce 3D models using the highest-quality settings in the field, we consistently took curation-quality photographs. We define curation-quality photographs as those capable of re-creating the object in the highest degree of fidelity, covering all aspects of its shape, documenting lighting conditions and camera settings, and that are shot in the RAW. This created the proper digital legacy for each project and allows us or future researchers to create high-resolution 3D models using the same data. Once we returned to the Mesoamerican Archaeology Lab at the University of California, San Diego, we were able to generate improved models that are adequate for structural analysis and digital curation. Our workflow included daily post-processing and archiving of digital models to prevent a common backlog of digital archaeological research (Forte et al. 2012;3; Myers and Badillo 2019; Olson et al. 2013).

Our models and resulting orthophotos were tied to the UTM coordinate system by using existing GCPs that were georeferenced with a Trimble GeoExplorer 6000 Series GeoXH GPS Unit with 10 cm accuracy (Braswell and Daniels 2013). With the 3D model open in Metashape Pro, we measured the horizontal distortion. This was based on the known position of our GCPs—markers placed with a tape measure and the corners of the unit. When distortion was within acceptable levels (5 cm in any direction), we were able to use the model for illustration (Figure 5a). The model was exported using the common OBJ file format, which can be used across several post-processing platforms. Antón and colleagues (2018) found that the model geometry across the mesh was consistent across the most popular file formats (PLY, OBJ, FBX, 3DM, and 3DS).

Illustration, Presentation, and Consolidation

We exported an orthomosaic, a scaled plan view of the model made from stitching images together, depicting the overhead view of the unit(s) as a TIFF file (Figures 5b and 5c). This file was opened in Adobe Illustrator and traced using the pen tool to create a scaled drawing of each unit across the structure (Figure 5d). We recognize the possibility of distortion when tracing...
the models. These illustrations are interpretations, based on a model accurate to a known measurement, that highlight relevant archaeological features. Unlike hand-drawn illustrations, our distortion is quantified and reported. The model is exported as a 3D PDF and uploaded to a tablet with a PDF-viewing app.

Architectural illustrations document the life history of the excavated structure, the different construction phases, the locations of caches and features, and any later modifications. They are simplified representations that show only the most important details. Yet when documenting features in the field, it is possible to miss or disregard artifacts that later prove to be significant (Douglass et al. 2015). The advantage of our method is that the orthomosaic documents everything in the scene. These overhead shots produced with a simple telescoping monopod (Figure 6a) provided an aerial view of the excavation equivalent to drone imagery (e.g., Brown et al. 2019; Howland et al. 2014), which we note require special permitting in Belize. When we create our simplified architectural illustration, we are able to make the decisions as to what to include in the lab, without time constraints. The illustrations included multiple layers that emphasized different kinds of information contained within the orthomosaic.

Both illustrations and models are used in the field during consolidation (Figure 6b). The TRIP ethos is to reconsolidate the building as a ruin. We use a dolomite-concrete-earth mix for mortar. The mortar is not as permanent as concrete, but it promotes better drainage and does not damage the soft limestone blocks as it dries. When dealing with this cultural heritage, our aim is not to re-create the buildings they may have been but to preserve the architecture in its current stage of life history. Each wall is only as many courses high as it was prior to starting excavations. Structure 6 yielded missing walls and incomplete wall segments (Braswell et al. 2019). The 3D visualization is invaluable during the reconstruction process.

Our models and illustrations provided a blueprint for those reconsolidations. This process not only reflects the life history of the building but ensures that the end result is of accurate proportion to the ruin. The models are scaled, and the final reconstructions are based on measurements from the model. The different layers of the illustration show the different architectural phases within the structure. We show these different phases in the final reconsolidation, using the illustration to guide our display of interior walls. This was particularly pertinent during the 2019...
excavations, in which we encountered several consecutive walls corresponding to different construction phases and buttress walls.

The monuments that we consolidate are the only experience some people have of the ancient Maya. There are up to 2,000 visitors a week at Nim Li Punit during peak season. The architectural illustrations are also a way of communicating the history of the structure to the academic community. For this reason, it is incredibly important to consolidate these features with fidelity to the original structure. In addition, the models provide a way for us to evaluate the idiosyncratic architecture of the southern Belize region. It is defined through a variety of markers, including tiered, or “wedding cake”-style, buildings, masonry tombs, and the incorporation of the natural hillside into the foundations of acropolises and large platforms (Braswell and Prufer 2009; Leventhal 1990, 1992). Nim Li Punit displays all of these traits (Borrero et al. 2016; Fauvelle et al. 2013). Yet, there are some irregularities with the construction of the excavated structures in the south and west groups.

Structure 6 was found to be missing several west-facing walls. Structure 50 had several buttressing walls against potentially

FIGURE 5. Example of orthomosaic used to produce an illustration for publication: (a) line drawing produced of F. 6/33W/1, a north-facing wall of the Structure 5 platform; (b) rotated view of Op. 6/33W/1, exposing fallen architecture and intact patio floor; (c) overhead orthomosaic of Op. 6/33W/1, in which can be seen the white board with metadata, corners of the unit tied into UTM grid, scale bar, north arrow, and GCPs; (d) detail of overhead drawing of Structure 6, including Op. 6/33W/1 in the southwest corner—an example of the scaled architectural drawings that can be made using this technique. (Made in Metashape and Adobe Illustrator.)
DISCUSSION

SfM provides several advantages as an efficient way of capturing provenience data of architectural features. The digital workflow has reduced hundreds of person-hours of field measurements, scanning, and tracing to short capturing sessions in the field and nightly processing sessions in the lab. Time saved in this way accelerated excavation. Another major advantage not afforded by traditional graph-paper drawings is the use of 3D models in the field—a valuable tool in interpreting and consolidating fallen architecture. Using the 3D models on a smartphone or tablet, we were able to visualize the architecture as it was pre-excavation, which allowed us to revisit artifacts or features removed during excavation. This was immensely helpful during reconsolidation, providing multiple views of the subject not afforded by hand-drawn representations. It enabled us to compare the models of unconsolidated structures from one season to another, finding similar patterns of missing or altered architecture.

In the broader world of cultural heritage, 3D models are being adopted into the Building Information Modeling as a design paradigm. This process seeks to replicate in the digital environment the actual physical structure and generate the associated data—such as color, texture, and volume—of each individual architectural element (Garagnani 2017). This allows us to analyze both construction history and building material. These data are important for future conservation efforts. Image capture and processing software has led to the possibility of building full-scale digital engineering models using low-cost digital cameras and personal computer software (Douglass et al. 2015; Khalaf et al. 2018; Pavlidis et al. 2006).

We are not the first to report the advantages of modeling architecture (Fritz et al. 2016; Howland et al. 2014; Porter et al. 2016), but we offer a way of independently evaluating its accuracy and demonstrate its applications for consolidating of architecture and comparing different excavated contexts. Our focus is on two sources of error: (1) the quality of image orientation when generating the sparse cloud, and (2) the horizontal distortion of the final model. The former can be compared across all models, regardless of scale; the latter is dependent on the subject, but relative percentages of distortion can still be measured (e.g., our 5 cm horizontal distortion is 2.5% of a single unit edge). Reporting error should be standard practice to allow peers to review models and determine their functionality for quantifiable analysis (Supplemental Table 1; Supplemental Figure 1). SfM photogrammetry in archaeology has matured over the last two decades, and there should be a common set of best practices.

There are some caveats. The first is that accuracy is only as good as the initial measurements. Our subjects were buildings several meters in length, so we placed our markers using tape measures. Our measurements are not accurate to the sub-centimeter level. This would pose a problem for the modeling of small artifacts, but it was acceptable to us for the purposes of documenting buildings. When we tie our model into the site map, there will be an acceptable level of error for the scale of the monumental core (see the above discussion of georeferencing the site). Higher levels of precision are achievable with a total station.

When considering photogrammetry in general, more disadvantages become apparent. There are issues of accessibility of the fallen-down walls. These can be independently visualized using our imaging techniques. Comparing the fall patterns between different structures facilitates anastylosis. While excavating Structure 50 in 2019, we were able to contrast the abundance of walls with the missing walls from the previous year. We concluded that Structure 6 had been deliberately disassembled in antiquity, whereas the buttress walls of Structure 50, combined with the bowed nature of the north-facing wall, indicated a response to architectural collapse.

The primary takeaway of the work presented here is that these visualizations, be they 3D models, 2D illustrations, or a consolidated ruin, are open to independent review based on reported errors. This is very important for archaeologists committed to transparency in scientific and public communication. We consider the consolidated ruins themselves to be a form of public outreach. We have demonstrated here that those consolidations are accurate to a known degree. Any ethical project that uses 3D modeling should begin with recording metadata and end with reporting paradata.
technology and the subject being digitized; workflow management and adapting it to changing conditions; data storage and dissemination, especially when working with a large number of photographs; and the lack of a set of best practices. Because there are various proprietary (e.g., ReCap, Metashape, Pix4D) and open-source (e.g., Bundler, MicMac, Meshroom, VisualSFM) programs available, it is difficult to troubleshoot and compare between different workflows. Accuracy and error, when reported, vary greatly—usually a function of the distance between GCPs (Fei Peng et al. 2017). This speaks again to the need for agreed-upon best practices. Another concern is the size of data. The number of photographs taken will vary for each project, but it can easily produce terabytes of data (Table 1). Dissemination of these models provides a unique challenge, particularly when considering long-term data storage (Fernandez 2019). What we have argued for in this article are potential solutions to some of these issues.

CONCLUSION

In our case study of applying SfM to the excavation and consolidation of architecture, we present two advancements. The first is in terms of work efficiency. The excavation of architecture requires the documentation of the structure pre- and post-excavation. During the 2018 field season at Nim Li Punit, we expedited this laborious process through close-range SfM photogrammetry. The 3D models aided in reconsolidation and ensured that, from a scientific communication standpoint, we were engaging in the accurate display of these ruins. The second major advance is the importance of reporting key elements that inform the interpretation of our models, including paradata, reprojection error, and horizontal distortion.

Affordable digital cameras and the abundance of SfM programs have made photogrammetry a viable alternative to traditional mapping techniques. It is likely to remain a staple of digital archaeology. For this reason, we argue that it is time to develop, as a community, an academic standard based on the best practices for the archaeological application of photogrammetry. After more than a decade of experimentation with photogrammetry and 3D modeling, there is still room for growth (Reilly 1991; Stuart 2015). We posit that the proper display and dissemination of paradata and error will assist in the creation of archaeological models that move beyond beautiful abstractions of real-world subjects, and they will convert the 3D data we produce into meaningful methods of preservation and recording that can assist in comparative, analytical, and reconstructive efforts. If 3D data is to be accessible and subject to independent review, as a community, we must engage in a dialogue about data storage, file formats for conservation, and the assessment of model validity (Fernandez 2019; Forte et al. 2012; Richards-Rissetto and von Schwerin 2017). RAW photography should be curated alongside the models whenever possible to allow for remodeling based on the original photographs. We suggest that model validity is best assessed based on quantifying error going in (quality of image orientation as measured by reprojection error) and error going out (distortion of distance between known GCPs). As stewards of the archaeological record, we should strive to preserve the material we encounter at the highest level of fidelity and with the greatest amount of confidence in the digital curation of our subjects. We mimic the call of transparency through paradata by Bentkowska-Kafel and colleagues (2012), acknowledging that our virtual efforts are presenting a particular interpretation of the past.

This is an issue that speaks to the importance of scientific communication. We have grounded our discussion of best practices in the exercise of consolidating architecture at Nim Li Punit. It is important to represent the structures we excavate in accurate ways, not only as represented in our publications, such as the illustrations, but also as the public engages with the architecture in the consolidated ruin. To ensure that at all levels these representations of the past are accurate, we open ourselves to independent review through the adoption of these aforementioned best practices.

Supplemental Materials

For supplemental material accompanying this article, visit https://doi.org/10.1017/aap.2020.11.

Supplemental Figure 1. Reporting RMS error in our models: (a) RMS error through time; (b) RMS error by time (in minutes) spent on each project.

Supplemental Table 1. Paradata Associated with the Project Models from the 2018 TRIP Field Season.

Acknowledgments

The authors are grateful to our advisor, Geoffrey E. Braswell, for his support during this project, and the feedback generated by the discussants in our 2019 SAA symposium on best practices in photogrammetry. Mario Borrero would also like to thank the UC Davis ARC for the support they provided and the 3D Photogrammetry for Cultural Heritage workshop they held. Our appreciation goes to the anonymous reviewers for Advances in Archaeological Practice for their insightful comments and suggestions on earlier versions of this article.

Data Availability Statement

This is the first time that our methods have been published. The relevant meta-paradata is included with this publication and the accompanying supplemental materials. All relevant data regarding the generation of the models from the 2018 field season are included in the text of the article. Display models are available to view and download on the project Sketchfab profile (sketchfab.com/TRIParchaeology). The photographs used to generate the models are contained on Toledo Regional Interaction Project hard drives.

NOTES

1. Lerma and Muir (2014) provide an excellent example of paradata description during the modeling process, whereas Havemann (2016) provides a frank discussion of the current state of paradata reporting. For a longer-form discussion on paradata and virtual heritage, we refer the reader to Bentkowska-Kafel and colleagues (2012).

2. RMS error takes the difference between an observed inputted value (typically GCPs) and the estimated value for user-defined points, squares it, finds the mean square value, and then finds the square root of that value. The equation to calculate it is RMS value = \sqrt{\sum(X^2 + Y^2)}^2, where Xs and Ys are the
REFERENCES CITED


Bentkowska-Kafel, Anna, Hugo Denard, and Drew Baker (editors) 2012 Paradata and Transparency in Virtual Heritage. Ashgate, Farnham, United Kingdom.


Irish, Mark David 2015 The Ceramics of Lubaantun: Stasis and Change in the Southern Belize Region during the Late and Terminal Classic. Master’s thesis, Department of Anthropology, University of California, San Diego.


AUTHOR INFORMATION

Mario Borrero (mborrero@ucsd.edu, corresponding author) and Luke R. Stroth
Department of Anthropology, University of California, San Diego, 9500 Gilman Drive, #0532, La Jolla, CA 92093-0532, USA