NB: This document presents a transcription of the published form of words given in the article below. This has been produced solely for an administrative exercise and should on no account be cited -please cite the actual article which can be accessed from the AJPA website:

BRIEF COMMUNICATION

Standard methods for creating digital skeletal models using structure-from-motion photogrammetry

Brianne Morgan, Andrew L. Ford, Martin J. Smith

Abstract

Objectives: This article assesses best practices for producing 3D digital cranial models through structure-from-motion (SfM) photogrammetry, and whether the metric accuracy and overall presentation of photogrammetric models are comparable to physical crania. It is intended to present a user-friendly standard method of creating accurate digital skeletal models using Agisoft PhotoScan.

Materials and methods: Approximately 200 photographs were taken of three different crania, and were separated into series consisting of 50, 75, 100, 150, and approximately 200 photos. Forty-five cranial models were created using different photo series and a variety of PhotoScan settings. These models were assessed based on defined qualitative criteria, and model measurement estimates were compared with physical skeletal measurements using Bland–Altman plots.

Results: The majority of all models (37/45) produced measurement estimates with mean differ- ences of 2 mm or less regardless of PhotoScan settings, and therefore demonstrated high levels of agreement with the physical measurements. Models created with 150 photographs and on “high” PhotoScan settings scored the highest in terms of qualitative appearance in the shortest amount of time.

Discussion: In PhotoScan, it is recommended to create cranial models using 150 photographs and “high” settings; this produces digital cranial models that are comparable to physical crania in both appearance and proportion. SfM photogrammetry is a convenient, noninvasive, and rapid 3D modeling tool that can be used in almost any setting to produce digital models, and following the guidelines established here will ensure that these models are metrically accurate.

1 | INTRODUCTION

In recent years, the idea of digital computer-based recording and anal- ysis of human remains has attracted increasing attention as relevant technologies have been developed and refined. This point has particu- lar currency in relation to the three-dimensional reconstruction of skeletal material. Digital skeletons have been proposed as a potential solution to the loss of data associated with repatriation and reburial, and offer a nondestructive, less intrusive, and more efficient method of capturing aspects of human skeletal variation (Kuzminsky & Gardiner, 2012). Some analyses are in fact only possible through the use of virtual skeletons, such as contour-based approaches for determining ancestry affiliation, or volume-based approaches for assessing sex, and many researchers argue that such methods provide a basis for analyzing skel- etal material in a more quantitative manner (Decker, Davy-Jow, Ford, & Hilbelink, 2011; Shearer, Sholts, Garvin, & Wärmländer, 2012; Sholts, Walker, Kuzminsky, Miller, & Wärmländer, 2011).

One popular method of creating digital skeletal models is structure-from-motion (SfM) photogrammetry, a technique wherein 3D structures are processed from overlapping 2D images (i.e., digital photographs) of the same object taken from a multi-angle, convergent perspectives using commercial off-the-shelf consumer-grade digital cameras (Morgan, Brogan, & Nelson, 2017). Since digital cameras can capture the appropriate level of detail required to generate models through SfM photogrammetry, it is more cost-effective than either laser scanning or CT imaging and its semi-automated methodology means that it is highly user-friendly (Micheletti, Chandler, & Lane, 2015). Its success in documenting and creating 3D models of multiple skeletons buried in mass grave environments suggests that SfM photogrammetry may also be ideal for capturing and digitizing individual skeletal elements as well (Baier & Rando, 2016).

A complete overview of the SfM photogrammetry workflow isbeyond the scope of this article, but in general, the process is broken down into four distinct stages. In the first stage, *sparse matching* is used to identify a relatively small number of features across multiple images, and assemble a relative sparse point cloud. This represents an initial model in 3D space, but more importantly the process also estab- lishes the initial *exterior orientation* (the 3D location and rotation of each *photo center*) Next, control/tie points (features with known 3D coordinates) may also be identified in one, some or all photographs, thereby giving the 3D model *scale* and allowing the *absolute* rather than *relative* exterior orientation. Dense-matching is then used to cre- ate a dense point cloud, usually on a pixel-wise basis, after which a mesh representing the surface topology will be reconstructed through 3D triangulation of these data points. The final step in generating a photogrammetric model is to add texture using the 2D photographs, which is accomplished by *orthorectifying* images to the corresponding polygons from the mesh (Fabio, 2001; Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012).

Validation studies have established the utility of photogrammetry in comparison to other methods of model creation, and shown that photogrammetric models produce comparable metric results (Evin et al., 2016; Katz & Friess, 2014). However, standard methods of model creation should be set out and validated, in order to assess the quality and texture of the models in comparison to the physical skele- tal elements, and to determine the level of osteological detail that photogrammetric models can provide. The objective of this study is to investigate best practices of photogrammetric model creation using Agisoft Photoscan, a popular SfM photogrammetry program, and to assess which settings and photography methods produce the best cranial models in terms of metric accuracy and appearance. Additionally, this study is intended to provide guidelines for photography setup and photogrammetric parameters that are both accessible (require minimal starting expertise) and workable (produces accurate skeletal models).

2 | MATERIALS AND METHODS

| Physical measurements

Three intact, adult crania were selected from Bournemouth University's Athenaeum Place collection (Figure 1a—c). Random points corresponding to small foramina or porosities were chosen and labeled on each crania; 20 on cranium 927 (labeled A—S, Z), 19 on cranium 674 (labeled A—S), and 18 on cranium 520 (labeled A—S, excluding K) were identified. As the goal of this project was to determine the accuracy of photogrammetric models (including textural resolution), small foramina or porosities were used as measurement points as opposed to standard anatomical land- marks. Using these points, 50 randomly selected measurements were taken with digital calipers to the nearest 100th of a millimeter on both crania 674 and 520, while 52 measurements were taken on cranium 927. Each measurement was taken three times, with 24 hr between repetitions, to account for intraobserver error. The mean and standard devia- tion were calculated for each set of repeated measurements.

The same photography set up and camera (a Canon E0S 500D with a resolution of 15.1 megapixels and a Canon EF-S 18–55 mm lens) was used for each cranium. The camera was placed on a Velbon DF 40 tri- pod, and the tripod was adjusted between three different heights (Figure 2a). For each of the tripod heights, the cranium was oriented in three different positions. Four white scales were taped to a white rotating platform, and each cranium was placed on this platform dur- ing photography (Figure 2b). The physical scale bars were chosen based on ease of visibility, though it was found that using different types of physical scale bars did not affect measurement accuracy as long as different points on the ruler were still identifiable in photo- graphs. A white cloth was placed underneath the rotating platform, and a white backsplash was placed around it, to ensure uniformity of the background. The camera was used on automatic mode, and the flash was turned on for all photos, because the fluorescent overhead lighting did not illuminate all features of the crania. A distance of approximately 50 cm was maintained between the tripod and the platform at all times, and the platform was rotated 15^{o} between each photograph, to a total of 360^{o} for each of the nine combinations of tripod position and cranium orientation. Approximately 200 photos were taken for each cranium. These photos were used to form photo series that aimed to provide full coverage of the cranium, consisting of 50, 75, 100, 150, and approximately 200 photos. All images were taken in JPEG format.

2.3 | Model creation

Agisoft PhotoScan Professional Edition version 1.2.4 was used to assem- ble the digital models, beginning with cranium 927. PhotoScan uses a four-step workflow (alignment, build dense point cloud, build mesh, and build texture), and each step has a number of adjustable settings that control the final model appearance, and the time it takes to generate each model. A computer with 32.0 GB of RAM and a 3.50 GHz processor was used for all analyses. Initially, models were assembled of cranium 927 using the highest *alignment accuracy* and ultra-high *dense point cloud quality* settings, while all other advanced settings for each step were left as the Agisoft defaults. Models for each photo set were generated on both high alignment and high point cloud quality, and then medium align- ment accuracy and medium dense point cloud quality. Based on initial tri- als, it was found that models produced on low settings performed poorly, so no further models were generated on low settings. Models are named based on the crania they represent, the number of photographs used to create them, and the settings they were created with (e.g., model 927-199-high was created from cranium 927 using 199 pictures on high alignment and high dense point cloud settings).

In PhotoScan, each of the 20 points picked from the physical cra- nium 927 were labeled directly on each of the 15 models using the “Create Marker” function. Scale bars (a tool in PhotoScan used to measure distances) between the markers that corresponded to the 52 physical measurements were created. These steps were repeated to generate 15 models each for cranium 674 and cranium 520.

2.4 | Model presentation evaluation

To determine their utility as a digital representation of a physical cranium, the models were evaluated on the criteria of (1) completeness; (2) texture; (3) edge definition; and (4) ease of marker placement. Each criterion was scored on a scale of 1–3, with 1 being the optimal grade, and 3 the least optimal. Figure 3 illustrates the scoring rubric for the first three criteria. For marker placement, a score of 1 was received if less than 20% of all markers were difficult to place, a score of 2 was assigned if 20–50% of markers were difficult to identify, and a score of 3 was assigned if more than 50% of the markers were difficult to place. Initially, every modelwas scored twice, and was rescored a third time in cases of disagreement between the first and second score. The final score was determined based on agreement between two of the three scores.

2.5 | Model measurement estimates

After model generation, PhotoScan Professional Edition allows users to scale the model using scale bars placed directly on the constituent photographs. In this case, scale bars were added to photos that con- tained rulers, and measurements were input to reflect the physical dimensions of the ruler. Three calibration scale bars measuring 1, 5, and 10 mm were added to each digital model on 5% of the photos used to produce each model. The accuracy settings were set to 0.001 m for all model measurements.

Measurement estimates were then generated. In order to assess the degree of randomness within the differences between the model esti- mates and the physical measurements, and to therefore confirm that they followed a normal distribution, frequency histograms were constructed for each model. Once it was confirmed that the differences in measure- ments were not systematic, the mean difference and the percent error between the measurement sets for each model were calculated. Bland– Altman plots were also constructed for each set of model measurement estimates. For each pair of model estimate and physical measurement, these plots compare the physical measurement (considered to be the “true” measurement) to the difference between the physical and model estimate (Bland & Altman, 1986). They visually represent the spread of the data by including the mean difference for all measurements pairs, and the upper and lower bounds of the 95% confidence interval of mean differ- ence (or, limits of agreement) (Bland & Altman, 1986). A Bland–Altman plot that displays narrow limits of agreement, a mean difference for all pairs of approximately 0, and data points that tend to cluster around the mean dif- ference represents two sets of data that agree (Bland & Altman, 1986).

3 | RESULTS

3.1 | Model presentation

The overall presentation of the models was highly variable, and depended both on the resolution settings and the number of photographs used to generate the model (Table 1). For all three crania, the model created with 150 pictures on high settings performed consistently well (Figure 4, Sup- porting Information Figures S1–S3). Certain features were more difficult to take photos of, which was reflected in the completeness of the models. For cranium 927, the left pterygoid process was incomplete on any of the models constructed with less than 150 photos, and for cranium 520, the internal lateral orbital surface was incomplete on all models. Many of the models, particularly those done on ultra-high quality settings, also displayed incompleteness of the internal nasal aperture.

Any differences in texture between models were found to be a result of the resolution settings, and using higher settings resulted in fewer blurry areas and better texture overall. Having a clear and well- defined texture also made it easier to place markers, since most of the markers chosen were small foramina or points of porosity. With many of the models generated using medium settings, the edges of the nasal aperture were overly smooth, and obscured many of the features that were present on the physical crania. On the other hand, many of the models generated on ultra-high settings displayed a jagged, non-continuous edge, which suggests that high settings are best for achieving smooth but realistic edges. The edges that scored the bestwere also associated with more photos.

3.2 | Measurementaccuracy

The frequency histograms produced for all models were unimodal and normally distributed, which means that the errors in measurements were random, and not related to systemic error in methodology. There- fore, mean difference and percent error were considered to be appro- priate representations of the accuracy of the data set. The majority of models produced measurement estimates with a mean difference of less than 2 mm, and a mean percent error of less than 2% (Table 2). Only 8 out of 45 models were outside these bounds, and most of these eight models were produced on medium or low settings with fewer than 100 photos. The models that scored the best during the model presentation assessment (927-150-high, 674-150-high, 520-150-high, 520-150-ultra high, and 520-195-high) all produced estimates that were well within the set accuracy parameters.

3.3 | Bland–Altman plots

Figure 5a–c represents the Bland–Altman plots of the highest scoring models in terms of appearance for each cranium. The Bland–Altman plot for model 927-75-low is also included, as an example of a model that does not display agreement between the physical measurements and the model estimates (Figure 5d). The Bland–Altman plots demonstrate that models created with more photographs and on higher settings tend to display better agreement between the digital measurement estimates and the physical measurements, with narrower limits of agreement and fewer or no data points outside of the 95% confidence interval limit.

| DISCUSSION

| Model presentation

The overall appearance of the virtual crania is initially the most impres- sive aspect of the photogrammetric models in this study. Many of the models represent complete crania that are well-textured and appear to accurately represent their physical counterparts, though models created with more photographs and on high settings performed the best overall. Other research analyzing the quality of cranial models created with SfM photogrammetry has found that noise tends to exist around particular areas, especially the foramen magnum, the orbitals, the zygomatic bones, and the parietals, which is consistent with the issues in resolution encountered in this study (Katz & Friess, 2014).

| Measurement accuracy

With the exception of eight models (most of which were made using medium or low settings), all models created in this study produced measurement estimates with high level of agreement to the physical measurements. The mean absolute difference for the best performing models ranged from 0.02 to 0.48 mm, and the majority of the mea- surement differences for all models are smaller than 2 mm, which is within the acceptable 2 mm error range allowed in osteometry. These values are consistent with accuracy results that have been obtained from measurements of CT and laser-generated 3D models, which suggests that they are just as appropriate for conducting osteometric analysis (Citardi et al., 2001; Dedouit et al., 2007; Fourie, Damstra, Gerrits, & Ren, 2011; Hildebolt, Vannier, & Knapp, 1990; Katz & Friess, 2014; Robinson et al., 2008; Stull, Tise, Ali, & Fowler, 2014; Verhoff et al., 2008).

4.3 | Recommendations for photography

The initial photography setup and execution is the probably the most important factor when assembling models using photogrammetry. If the photographs are of poor quality, the PhotoScan program will not be able to assemble a complete or accurate model. Any camera used for SfM photogrammetry should have a minimum resolution of 8 megapixels, and should be able to take photographs in either RAW or high quality JPG format (Costa Moraes, Dias, & Melani, 2014; Mallison & Wings, 2014). While this study and several others have had success in using JPEG images to create accurate photogrammetric models, it is still recommended to have nonlossy images available in SfM photogrammetry (Costa Moraes et al., 2014).

When photographing objects for the purpose of photogrammetry, it is also important to have as few shadows as possible in the images, as this can affect the alignment phase of model generation (Micheletti et al., 2015). For this reason, it is often recommended to only use ambient lighting sources, and to keep the camera flash turned off during the photography stage (Micheletti et al., 2015). In the current study, flash photography was used, as the ambient lighting sources were not bright enough to fully illuminate the texture of the cranium.

Positioning the tripod 50 cm away from the cranium and utilizing an all-white background to enclose the turntable area helped to ensure the use of flash would not create shadows and lighting errors, and none of the models appeared to be negatively affected by the use of flash. All of the crania used in this project were also highly textured and were not reflective, which also helped to reduce any camera flash-related issues, as many problems that arise from poor lighting are associated with shiny surfaces (Micheletti et al., 2015). Depending on the cranium being modeled, in future applications it is recommended to have a better source of ambient lighting so that the flash can be turned off, which will eliminate the potential for lighting errors.

During the photography phase of this study, photos were taken using the camera's automatic mode, which should have allowed it to remain on optimal settings for the duration of the session. This approach worked well for most camera angles, although it sometimes created problems with the depth of field, and it was often challenging to consistently keep all cranial features in focus. Ideally, a good photography and camera setup will use f-stop settings in the mid-to-high range and the "aperture-priority" mode to maximize the depth of field.

Aside from ensuring that camera and settings are appropriate for capturing images for SfM photogrammetry, it is also important to ensure that the setup and method of photography is conducive to high quality photos. In this study, a rotating platform method was used, and is recommended for creating photogrammetric models of small objects (Mallison & Wings, 2014). When using the turntable method, it is important to ensure that the background (the section of the photographs that is not rotating) is blank, so that the photogrammetry program does not attempt to align the still features in the background at the expense of the moving features in the foreground (Mallison & Wings, 2014). The white background used in this setup also likely contributed to the efficacy of the photographs; the background should be similar in color to the object being photographed, as high contrasts can create shadows and distortions.

Based on the results obtained in this study, it is obvious that the number of photos used to generate the model will have a significant impact on the quality of the final product. Previous studies that have used SfM photogrammetry to reconstruct skulls have used anywhere from 65 to 120 photos, so there is clearly a wide range of pictures that can result in a valid digital representation (Costa Moraes et al., 2014; Katz & Friess, 2014). While using as few as 50 pictures was sufficient to generate a complete cranium, the models created with this number of pictures were generally of poorer quality. Optimizing redundancy by including more pictures not only ensures that every feature of the skull can be captured, but it also allows for more leniency in the quality of photos that are used (Westoby et al., 2012). However, the relationship between time and number of photographs means that it is practical to limit the number of photographs in order to allow a reasonable timeframe for model generation. The balance between model quality and generation time was met at around 150 photographs in this study; beyond this number, there was no visible improvement in quality or measurement agreement, but the time it took to create a model increased significantly. It is, therefore, recommended that 150 pictures be used to create the optimal cranial models using Agisoft PhotoScan (Table 3).

4.4 | Recommendations for model creation settings

None of the PhotoScan resolution settings that were manipulated in this study had a discernible impact on model measurement accuracy, and models that achieved high levels of agreement between physical measurements and model estimates were created on all resolution settings. However, the best looking models in the shortest time frame were consistently produced on high settings. Therefore, for producing basic, but good quality models in Agisoft PhotoScan, high alignment accuracy and high dense point cloud quality settings are recommended (Table 3). While only manipulating some of the many settings and tools PhotoScan has to offer can limit the final model output, the program was still able to produce metrically accurate cranial models using mostly default settings and with minimal effort and user expertise, therefore fulfilling the objectives of the study.

ACKNOWLEDGMENTS

This work was undertaken as part of a broader initiative to explore the potential of photogrammetry in application to human remains at Bournemouth University through the Digital REconstruction And Measurement of the Skeleton (DREAMS) project. We would also like to thank our reviewers for their valuable feedback.

REFERENCES

Baier, W., & Rando, C. (2016). Developing the use of structure-from motion in mass grave documentation. Forensic Science International, 261, 19–25.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet, 327 (8476), 307–310.

Citardi, M. J., Herrmann, B., Hollenbeak, C. S., Stack, B. C., Cooper, M., & Bucholz, R. D. (2001). Comparison of scientific calipers and computer enabled CT review for the measurement of skull base and craniomaxillofacial dimensions. Skull Base, 11(01), 5–12.

Costa Moraes, C. A.,Dias, P. E.M., & Melani, R. F. H. (2014). Demonstration of protocol for computer-aided forensic facial reconstruction with free software and photogrammetry. Journal of Research in Dentistry, 2(1), 77–90.

Decker, S. J., Davy-Jow, S. L., Ford, J. M., & Hilbelink, D. R. (2011). Virtual determination of sex: Metric and nonmetric traits of the adult pelvis from 3D computed tomography models. Journal of Forensic Sciences, 56(5), 1107–1114.

Dedouit, F., Telmon, N., Costagliola, R., Otal, P., Joffre, F., & Rougé, D. (2007). Virtual anthropology and forensic identification: Report of one case. Forensic Science International, 173(2), 182–187.

Evin, A., Souter, T., Hulme-Beaman, A., Ameen, C., Allen, R., Viacava, P., Dobney, K. (2016). The use of close-range photogrammetry in zooarchaeology: Creating accurate 3D models of wolf crania to study dog domestication. Journal of Archaeological Science: Reports, 9, 87–93.

Fabio, R. (2001). From point cloud to surface: The modeling and visualisation problem. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34, 5–10.

Fourie, Z., Damstra, J., Gerrits, P. O., & Ren, Y. (2011). Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. Forensic Science International, 207(1), 127–134.

Hildebolt, C. F., Vannier, M. W., & Knapp, R. H. (1990). Validation study of skull three-dimensional computerized tomography measurements. American Journal of Physical Anthropology, 82(3), 283–294.

Katz, D., & Friess, M. (2014). 3D from standard digital photography of human crania—A preliminary assessment. American Journal of Physical Anthropology, 154(1), 152–158.

Kuzminsky, S. C., & Gardiner, M. S. (2012). Three-dimensional laser scanning: Potential uses for museum conservation and scientific research. Journal of Archaeological Science, 39(8), 2744–2751.

Mallison, H., & Wings, O. (2014). Photogrammetry in paleontology—A practical guide. Journal of Paleontological Techniques, 12, 1–31.

Micheletti, N., Chandler, J. H., & Lane, S. N. (2015). Structure from motion (SfM) photogrammetry. In S. J. Cook, L. E. Clarke, & J. M. Nield (Eds.), Geomorphological techniques (Chap. 2, Sec. 2.2.). London, UK: British Society for Geomorphology.

Morgan, J. A., Brogan, D. J., & Nelson, P. A. (2017). Application of structure-from-motion photogrammetry in laboratory flumes. Geomorphology, 276, 125–143.

Robinson, C., Eisma, R., Morgan, B., Jeffery, A., Graham, E. A., Black, S., & Rutty, G. N. (2008). Anthropological measurement of lower limb and foot bones using multi-detector computed tomography. Journal of Forensic Sciences, 53(6), 1289–1295.

Shearer, B. M., Sholts, S. B., Garvin, H. M., & Wärmländer, S. K. (2012). Sexual dimorphism in human browridge volume measured from 3D models of dry crania: A new digital morphometrics approach. Forensic Science International, 222(1), 400e1–400e5.

Sholts, S. B., Walker, P. L., Kuzminsky, S. C., Miller, K. W., & Wärmländer, S. K. (2011). Identification of group affinity from crosssectional contours of the human midfacial skeleton using digital morphometrics and 3D laser scanning technology. Journal of Forensic Sciences, 56(2), 333–338.

Stull, K. E., Tise, M. L., Ali, Z., & Fowler, D. R. (2014). Accuracy and reliability of measurements obtained from computed tomography 3D volume rendered images. Forensic Science International, 238, 133–140.

Verhoff, M. A., Ramsthaler, F., Krähahn, J., Deml, U., Gille, R. J., Grabherr, S., Kreutz, K. (2008). Digital forensic osteology—Possibilities in cooperation with the Virtopsy® project. Forensic Science International, 174(2), 152–156.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300–314.

How to cite this article: Morgan B, Ford A, Smith M. Standard methods for creating digital skeletal models using structure-from-motion photogrammetry. Am J Phys Anthropol. 2019;1–9. https://doi.org/10.1002/ajpa.23803