Structure-from-Motion Photogrammetry in Earth Sciences and Paleontology

Structure-from-Motion (SfM) photogrammetry is now widely applied in earth sciences and paleontology, both as a research support and pedagogical tool (Breithaupt and Matthews, 2001; Breithaupt et al., 2001; Fahlke and Autenrieth, 2016; Matthews et al., 2016). This technique allows to merge sets of overlapping images taken with parallax angles and uses stereo-matching to generate 3D models (Granshaw, 2016; Brecko and Mathys, 2020). In geosciences, such models are used for samples of diverse nature -molds, traces, minerals, fossils, outcrops - covering a broad range of sizes, geometries, and textures. The resulting reconstructions can range from centimeter-scale specimens (Bennett et al., 2021; Matthews et al., 2016; Falkingham et al., 2018) to landscapes spanning several tens of meters (Voulodimos et al., 2016; Breithaupt et al., 2021).

Once created, these models can be archived and shared on platforms such as Sketchfab (http://www.sketchfab.com), which allow interactive 3D visualization. Users can explore models at various magnifications, examine morpho-structural details, change perspectives, and compare the digital model to the original sample for teaching or research purposes. For example, the French Geological Society's virtual atlas (https://www.geosoc.fr/geologie-virtuelle.html) already hosts about 850 models, regularly enriched by university contributors.

A variety of acquisition and processing methods are available. We developed a specific workflow that balances geometric accuracy with high-quality surface texturing. The setup requires DSLR cameras or equivalent hybrid system, a controlled-light photographic studio, a turntable, and a custom rig to capture overlapping views with parallax.

Our study focuses on objects ranging from 2 to 40 cm in size and up to 5 kg in weight, covering a broad spectrum of fossils and geological samples.

Image Acquisition

The capture design uses a convergent setup with four synchronized cameras mounted on an adjustable, custom-built rig arranged in an arc-circle shape (Figure 1a). Samples are placed on an automated turntable (Foldio 360© Orange Monkey), that triggers all 4 cameras synchronously via infrared. A rotation-step is 15°, raising 96 pictures for a full 360° rotation of a single sample position. To record the full surface, the sample is repositioned after each sweep, with two to four positions typically required. Depending on surface complexity, the step size may be reduced to 10° and additional orientations added, resulting in a number of pictures ranging from 192 to 576 per specimen.

We use Canon EOS 77D DSLR (APS-C, 24.2 MP, 6000 × 4000 pixels, 22.3 × 14.9 mm CMOS sensor) cameras with two interchangeable macro lenses: a 60 mm f/2.8 for larger objects, and a 100 mm f/2.8 for smaller or more detailed ones. A circular polarizing filter used for reflective or translucent specimens reduces glare without altering color fidelity. Depth of field is managed by carefully balancing camera-object distance and diaphragm aperture, with focus set at one-third of the object depth before switching to manual mode. To preserve detail in highlights and reduce motion blur, images are slightly underexposed and triggered at high shutter speed.

To ensure accurate scaling and correct lens distortion, samples are placed on a customized circular target with embedded coded markers, adapted from Porter et al. (2016). The size of this background image can be rescaled to fit different specimen dimensions, while the surface occupation should be kept around 15-25% of the total area, for keeping textured pattern and markers always visible.

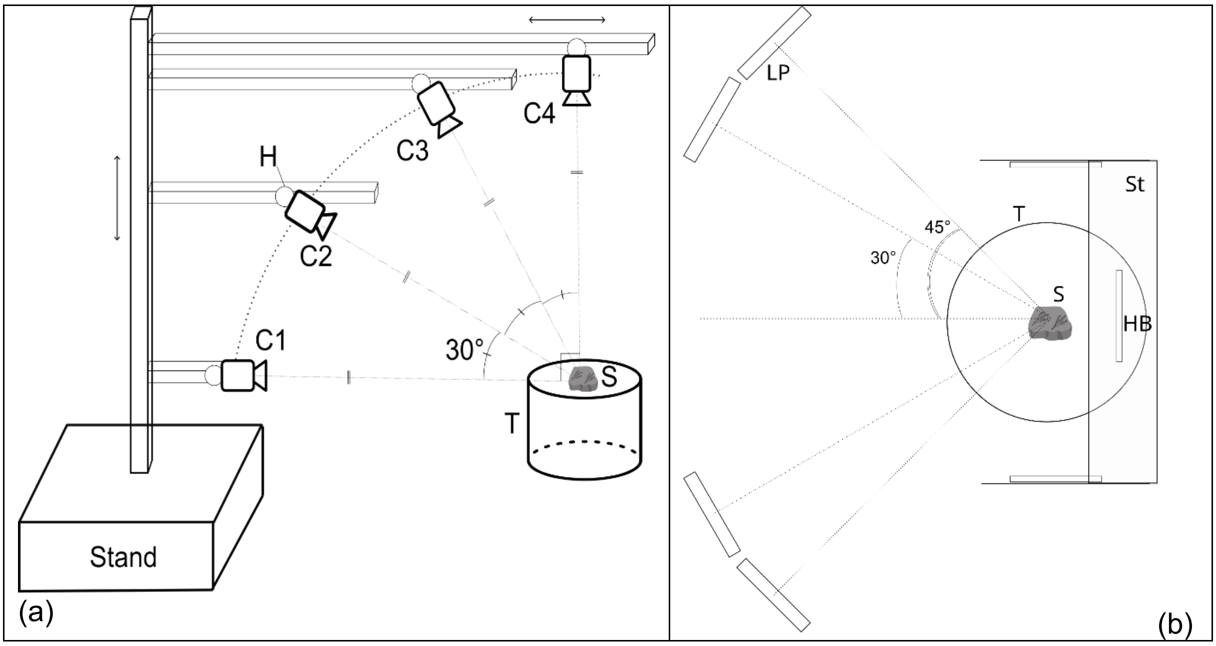

Figure 1 : (a) - Photographic setup. C1 to C4 : synchronized cameras at equal distance to the sample S. T : turntable with default 15° rotation step and infrared triggering device. H : Ball Head. Horizontal arms on the stand and camera ball heads can be displaced vertically and horizontally. (b) - Plan view of the lighting configuration. St: Studio box; HB: Halo bar; LP: Led panel.

Lighting (Figure 1b) are controlled with a combination of four dimmable LED panels (Dorr 3040BC©) and a Foldio 360© light-box fitted with LED halo bars and dimmers, ensuring constant, diffuse daylight-equivalent illumination (5600 K). The light box is set with velvet backdrops to ensure background contrast with the specimen. LED panels are placed at ±30° and ±45° relative to the sample to minimize shadows and reflections.

Photographs are captured in 14-bit RAW format. A ColorChecker© chart is included in each session for color calibration, white balance, and post-processing consistency. Calibration is performed with ColorChecker Camera Calibration© software, generating ICC color profiles applied in Capture One©. Batch processing helps ensuring uniform adjustments across all images. Final exports are saved in full quality .jpeg format to balance image sharpness and storage efficiency.

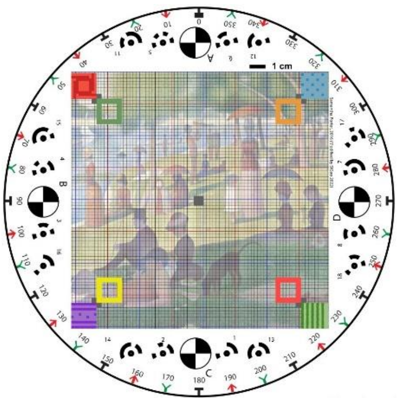

Figure 2 : Background image with textured pattern and coded markers, modified from Porter et al. (2016). It is set down on the turntable with the sample.

Photogrammetric Processing

The computing hardware comprises a high-end workstation with Graphic Processing Unit. The photogrammetry software is Agisoft Metashape Professional© v1.8, chosen for its robustness, user-friendly interface and efficiency. A summary of the hardware configuration and a typical project size is given in table 1.

Processor	2 Intel ® Xeon® Silver 4114 CPU @2.20GHz
Graphic card	NVidia GeForce GTX 1080, 8Gb RAM, 2560 Cuda cores
RAM	64 Gb
Images per specimen	192-576
Disk volume of RAW images	5-17 Gb
Disk volume of JPG images	3-10 Gb
Filtered Sparse point cloud count	~100k pts
Filtered Dense point cloud count	~2M pts
Final mesh	100-500k faces
Full PSX file volume	1-5Gb
Exported OBJ file volume	<200Mb

Table 1 : Hardware configuration and size of typical 3D projects

The detailed processing workflow is mainly similar to the one of Over (2021) and uses the same parameters for filtering and error reduction, but differs in some parts from the fact that their acquisition is georeferenced while our is only scaled. Our spatial reference is metric local coordinates, and our main end-product is the textured mesh in an arbitrary geometry i.e. full 3D geometry. The table 2 below shows the main differences from Over (2021). Each sample is captured in several positions, providing a separate photogrammetry project called a chunk. For each chunk, we make the same workflow until the calculation of a first mesh, allowing a mask from model to be back-propagated on the pictures. Then, all the pictures from the different chunks are merged into a single one, keeping on the picture the masked area previously determined, and the photogrammetry process is fully re-run, using the scales to optimize scene geometry and to set a metric scale onto the 3D model. This scaling occurs in the optimization step. More detail can be found on the original workflow designed by Over (2021).

Alignment	
Exclude stationary points	Yes
Reference settings	
Maker accuracy	0.0005m
Tie point accuracy	0.1 pix
Build Model	
Source data	Depth maps
Surface type	Arbitrary
Quality	Medium
Filtering	Mild
Interpolation	Disabled
Build Texture	
Texture type	Diffuse map
Source data	Images
Mapping mode	Generic
Blending mode	Mosaic
Texture size	8192

Table 2 : Agisoft Metashape’s processing parameters specific to our study

SfM-photogrammetry typically follows four stages described below:

Preprocessing - A quality check on the images is performed, and all suitable images are imported in the project workspace. A mask is created by an automatic procedure on every picture to exclude all parts around the sample and its background target (Figure 2). Excluding stationary points during alignment is set on in order to cope with errors in mask limits.
Camera alignment – This step involved detecting feature points with SIFT algorithm across images, then matching tie points. The camera positions are computed via bundle-block adjustment, then generating a sparse point cloud from each tie point. Point cloud quality and 3D scene optimization is refined by filtering reconstruction uncertainty, projection accuracy, and reprojection error. Scaling and camera calibration are achieved using the pre-defined markers on the acquisition target image (Figure 2), with a root mean square error (RMSE) threshold set to 0.1 mm.
Depths map generation and dense cloud – Depths maps are calculated using cross-correlation stereo matching on sub-sampled image pairs. They are then merged into dense clouds, with mild filtering to remove inaccurate points while keeping micro-morphologies. Medium-quality settings are sufficient to balance detail with manageable processing times.
Meshing and Texturing – Model reconstruction used interpolation on the dense cloud and generated a 3D triangular mesh. Note that in new version of Agisoft Metashape >2.0, model can be calculated directly from the depth maps without the need of producing a dense cloud. Mesh holes and outliers are corrected, and optional mesh refinement or smoothing applied. A connectivity filter is necessary to ensure that only the main sample triangular faces are kept without floating pieces. Final UV mapping projects a selection of high-quality images onto the mesh. Images capturing critical morphological details or overlap areas between chunks are prioritized, while low-quality inputs are masked out. Mosaic blending minimizes seamlines artifacts.
The finished textured meshes are exported in .obj format and uploaded to Sketchfab, along with metadata, but sometimes after decimation because our professional account only affords for a maximum project size of 200 Mb. Complete Metashape project files, raw image sets, and metadata are systematically archived in our institutional repository for long-term preservation and reuse.

References

Bennet MJ. 2021. On the nature of extreme close-range Photogrammetry : visualization and measurement of North African stone point. *Code(4)lib Journal* 51:2021-06-14.

Brecko J, Mathys A. 2020. Handbook of best practice and standards for 2D+ and 3D imaging of natural history collections. *European Journal of Taxonomy* 623. https://doi.org/10.5852/ejt.2020.623.

Breipthaupt BH, Chan MA, Seiler WM, Matthews NA. 2021. Wearthering pits versus tample marks: a re-interpretation of the « Dinosaur dance floor » : a Jurassic Navajo sandstone surface in the Vermilion Cliffs National Monument, Arizona. *Palaios* 36:331–338. https://doi.org/10.2110/palo.2020.077

Breipthaupt BH, Matthews NA. 2001. Preserving paleontological ressources using photogrammetry and geographic information systems In D Harmon (Ed.), Crossing Boundaries in Park Management: Proceedings of the 11th Conference on Research and Resource Management in Parks and on Public Lands. *The George Wright Society*:62-70

Breipthaupt BH, Matthews NA, Noble TA. 2004. An integrated approach to three-dimensional data dollection at dinosaur tracksites in the Rocky Mountain West. *Ichnos* 11:11–26. https://doi.org/10.1080/10420940490442296

Fahlke JM, Autenrieth M. 2016. Photogrammetry versus Micro-CT scanning for 3-D surface generation of a typical vertebrate fossil- a case study. *Journal of paleontological techniques* 14:1-18.

Falkingham PL, Bates KT, Avanzini M, Bennett M, Bordy EM, Breithaupt BH, Castanera D, Citton P, Díaz-Martínez I, Farlow JO, Fiorillo AR, Gatesy SM, Getty P, Hatala KG, Hornung JJ, Hyatt JA, Klein H, Lallensack JN, Martin AJ, Marty D, Matthews NA, Meyer CA, Milàn J, Minter NJ, Razzolini NL, Romilio A, Salisbury SW, Sciscio L, Tanaka I, Wiseman ALA, Xing LD, Belvedere M. 2018. A standard protocol for documenting modern and fossil ichnological data. *Paleontology* 61: 469–480. https://doi.org/10.1111/pala.12373

Granshaw SI. 2016. Photogrammetric terminology: third edition. *The Photogrammetric Record* 31: 210–252. https://doi.org/10.1111/phor.12146

Matthews NA, Noble T, Breiopthaupt BH. 2016. Close range photogrtammetry for 3-D ichnology: the basics of photogrammetric ichnology. In : Falkingham PL, Marty D, Richter A, Eds. Dinosaur tracks: the next steps. Bloomington (IND, USA) : Indiana University Press, 2016, pp. 28–55.

Over JR, Ritchie AC, Kranenburg CJ, Brown JA, Buscombe D, Noble T, Sherwood CR, Warrick JA, Wernette PA. 2021. Processing coastal imagery with Agisoft Metashape Professional Edition, version 1.6—Structure from motion workflow documentation: *U.S. Geological Survey Open-File Report*:2021–1039. https://doi.org/10.3133/ofr20211039

Porter ST, Roussel M, Soressi M. 2016. A simple photogrammetry rig for the reliable creation of 3D artifact models in the field: lithic examples from the early Upper Paleolithic sequence of Les Cottés (France). *Advances in Archaeological Practice* 4:71–86. https://doi.org/10.7183/2326-3768.4.1.71

Voulodimos A, Doulamis N, Fritsch D, Makantasis K, Doulamis A, Klein M. 2016. Four dimensional reconstruction of cultural heritage sites based on photogrammetry and clustering. *Journal of Electronic Imaging* 26:011013. https://doi.org/10.1117/1.JEI.26.1.011013