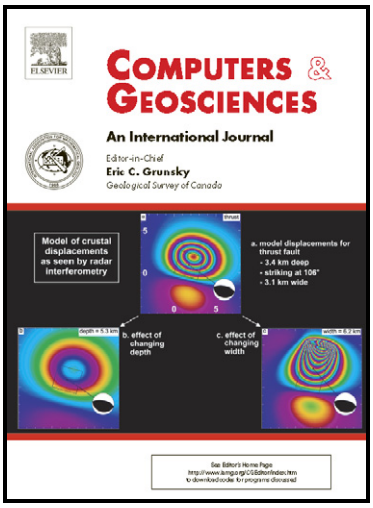
Author's Accepted Manuscript



www.elsevier.com/locate/cageo

Multiview 3D reconstruction in geosciences

M. Favalli, A. Fornaciai, I. Isola, S. Tarquini, L. Nannipieri

PII: S0098-3004(11)00312-8
DOI: doi:10.1016/j.cageo.2011.09.012
Reference: CAGEO 2703

To appear in: *Computers & Geosciences*

Received date: 22 June 2011
Revised date: 18 August 2011
Accepted date: 19 September 2011

Cite this article as: M. Favalli, A. Fornaciai, I. Isola, S. Tarquini and L. Nannipieri, Multiview 3D reconstruction in geosciences, *Computers & Geosciences*, doi:10.1016/j.cageo.2011.09.012

Multiview 3D reconstruction in geosciences

M. Favalli *, A. Fornaciai, I. Isola, S. Tarquini, L. Nannipieri

Istituto Nazionale di Geofisica e Vulcanologia, via della Faggiola 32, 56126 Pisa, Italy

Received 22 June 2011
Received in revised form 18 August 2011
Accepted 19 September 2011

Keywords:
Multiview
3D reconstruction
Laser scanner
Outcrop

* Corresponding author.
E-mail address: favalli@pi.ingv.it (M. Favalli).

ABSTRACT

Multiview three-dimensional (3D) reconstruction is a technology that allows the creation of 3D models of a given scenario from a series of overlapping pictures taken by using consumer-grade digital cameras. This type of 3D reconstruction is facilitated by freely available software which does not require expert-level skills. This technology provides a 3D working environment which integrates sample/field data visualization and measurements tools. In this study, we test the potential of this method for 3D reconstruction of decimeter-scale objects of geological interest. We generated 3D models of three different outcrops exposed in a marble quarry and two solids: a volcanic bomb and a stalagmite. Comparison of the models obtained in this study using the presented method with those obtained by using a precise laser scanner shows that multiview 3D

reconstruction yields models that present a root mean square error/average linear dimensions between 0.11 and 0.68%. Thus this technology turns out to be an extremely promising tool which can be fruitfully applied in geosciences.

<h1>1. Introduction

Multiview 3D reconstruction is the computationally complex process by which a full 3D model of a target scene is derived from a series of overlapping pictures of the target itself. The method lies at the frontier of computer vision research, and relies also on older methods used in photogrammetry (Mikhail et al., 2001). The large distribution of high-resolution (≥ 10 megapixels) consumer-grade cameras and the free availability of open-source programs implementing structure from motion (SfM) methods make 3D reconstruction from multiview simple and low cost. SfM is a process used to estimate both the scene geometry and the camera parameters (Hartley and Zisserman, 2004). Intrinsic camera parameters are either known a priori (Nister, 2004) or recovered a posteriori through autocalibration (Triggs, 2000). In a typical SfM procedure, the first step is the identification of distinctive features (key points) in the input images. Then a bundle adjustment algorithm allows the reconstruction of the 3D geometry of the scene by optimizing the 3D location of key points, the location/orientation of the camera, and its intrinsic parameters (Lourakis and Argyros, 2008; Triggs et al., 2000).

Recently, point clouds produced by bundle adjustment methods have been widely used to create models of architectural and bare earth surfaces with high accuracy (de Matías et al., 2009; Dowling et al., 2009; Grzeszczuk et al., 2009). Similar high-resolution 3D photorealistic models of geological outcrops constitute virtual outcrops which are ideal for visualization and quantification of 3D structural or sedimentary features, maximizing the benefit of field excursions (Bellian et al., 2005; McCaffrey et al., 2005; Pringle et al., 2006; Buckley et al., 2008). Photogrammetric surveys and computer vision techniques have been also used by James et al. (2007) to characterize morphological modifications of an advancing lava flow.

In this study we evaluate the performance of the multiview 3D reconstruction method in geosciences. We create 3D models of three outcrops exposed on a marble quarry and two solid samples of geological interest: a volcanic bomb and a stalagmite. Our examples

have typical linear dimension up to 1 m. For this purpose we defined a sequence of automatic steps which only uses freely available software and does not require any prior information on camera position, orientation, or internal camera parameters. The accuracy of the obtained models is assessed by comparison with models obtained by using a laser scanning technology.

<h1>2. Methods

<h2>2.1. Multiview 3D reconstruction

Multiview 3D reconstruction creates a 3D model starting from a series of overlapping photos imaging a given scene. This is achieved by running a series of algorithms which work automatically without a priori specification of parameters for the input pictures. The procedure applied in this work comprises the following steps: (i) the scale invariant feature transform (SIFT) algorithm (Lowe, 2004) is used for key-point extraction; (ii) the open-source SfM software package Bundler (Snavely et al., 2006, 2007) generates a sparse 3D point cloud with internally consistent 3D geometry; (iii) the open-source PMVS2 (Patch-based Multiview Stereo software – version 2) software takes the output of the Bundler software as input to reconstruct the model of the imaged scene in the form of a denser point cloud (Furukawa and Ponce, 2007, 2009); (iv) additional software is used for visualization and postprocessing. In the following sections, more details are provided for each step.

<h3>2.1.1. Recommendations for photoacquisition

The sequence of pictures, which constitute the starting input, must be taken from several viewpoints which vary significantly from one another. As an example, many pictures from the same viewpoint are useless, while pictures taken at each step by moving around the scene of interest are ideal. The sequence of pictures must be acquired while the target scene/object is fixed in the same position under a good lighting, and moving shadows and/or camera flash should be avoided as much as possible. In addition, the color texture (nonhomogeneity) of the object/scene of interest is important, because the procedure works on color changes. The theoretical minimum number of input photos is 3, but a minimum of 4 to 6 pictures is recommended to obtain reliable models, and the

model accuracy increases if a much higher number of "good" pictures is used (from tens to hundreds of pictures). An example of a good sequence of viewpoints (one picture from each viewpoint) is given in Fig. 1.

<h3>2.1.2. Feature extraction

In the first step of the global procedure, all the pictures are processed in loop by a pattern recognition algorithm and matched to each other to find corresponding features in different images. In this way a series of key points is obtained. This process is carried out by using the scale invariant feature transform (SIFT) algorithm (Lowe, 2004). A demo version of SIFT is available at http://www.cs.ubc.ca/~lowe/keypoints/. This demo (at present) works only on small images (a few megapixels); hence we down-sampled input images to fulfill this constraint. To simplify the processing, the input pictures are converted into gray scale images before running the SIFT algorithm. Regions of interest are those marked by sharp gradients in gray values. Typically, SIFT will detect up to tens of thousands of such features in a resampled image.

<H3>2.1.3. Structure from motion processing

Once corresponding key points have been identified across a series of images, the change in key-point position in different images is considered in the SfM process to clump the position of such points in a 3D reference system. This complex process takes also into account the focal length and sensor width of the camera used to take the image (the camera type is tagged in the header of the picture file). The output provides camera parameters and position for each considered input image by using a numeric optimization technique called ''bundle adjustment.'' In this work we use Bundler software (http://phototour.cs.washington.edu/bundler) which is an open-source SfM software (Snavely et al., 2006) that iteratively considers an increasing number of input pictures, providing an increasingly optimized output as the process goes on. If an input image is "not good" (e.g., it is blurred) Bundler automatically discards it. Bundler outputs also a sparse cloud of 3D points representing the imaged scene.

<H3>2.1.4. Dense 3D point cloud reconstruction

The output obtained from Bundler is then processed by the Patch-based Multi-View Stereo – version 2 package (PMVS2, Furukawa and Ponce, 2007; 2009). An open-source implementation of PMVS2 is available at http://grail.cs.washington.edu/software/pmvs/.

This further processing produces a much denser 3D point cloud which provides a very detailed and realistic model of the imaged scene. One of PMVS2 advantages is that it preserves only rigid structures (e.g., pedestrians walking in front of a monument will not be seen in the final result). PMVS2 is also robust against differences in image colors due to exposure settings, white balance, or lighting conditions. Various parameters and flags can be specified in the PMVS2 option file including the subsampling rate of images before the processing; a tentative density of reconstruction; the minimum number of images in which a point must be visible to be reconstructed; the minimum photometric consistency measure necessary to keep a point in the reconstruction (for details, see http://grail.cs.washington.edu/software/pmvs/documentation.html).

<H3>2.1.5. Visualization, surface reconstruction, and postprocessing

In the case of simple and substantially flat geometries, as for most outcrops, postprocessing can be done in a GIS environment by treating the obtained 3D models as digital elevation models (DEMs) with x and y coordinates assigned along the plane fitting the sampled surface and the elevation set orthogonally to this plane.

In the case of more complex 3D geometries, such as the two solid samples, a series of freely available tools have been used for the postprocessing, the rendering, and the error assessment of the obtained 3D models. 3D point clouds have been managed using the Scanalyze software, developed by the Stanford Computer Graphics Laboratory, freely available at http://graphics.stanford.edu/software/scanalyze/. Scanalyze is a computer graphics program for viewing, editing, and merging range images to produce denser polygon meshes (Besl and McKay, 1992; Levoy et al., 2000). The open-source MeshLab software has been used to connect the points cloud generated by PMVS2 in a network of triangles which approximates the continuous surface of the imaged scene. MeshLab allows the editing of unstructured 3D triangular meshes. This freely available software has been developed by the Visual Computing Lab of ISTI-CNR in Pisa, Italy (http://meshlab.sourceforge.net/). Finally, to compare the difference between pairs of complex surfaces we have used Metro, a tool designated to evaluate the difference between two triangular meshes (Cignoni et al., 1998). The mean distance E_m of a surface S1 from a surface S2 is defined as the surface integral of the distance divided by the area of surface S_1,

$$E_m(S_1, S_2) = \frac{1}{|S_1|} \int_{S_1} e(p, S_2) ds \,, \quad (1)$$

where $e(p, S_2)$ is the distance between a point p (belonging to S_1) and the surface S_2. Indeed Metro compares two triangular meshes S_1 and S_2 numerically (see Cignoni et al., 1998, for further details).

<h2>2.2. Laser scanning reconstruction

3D control models of the selected test surfaces have been obtained by using Konica Minolta VI-910 laser scanning, a noncontact 3D digitizer (www.konicaminolta-3d.com). The target surface is scanned by a laser beam (wavelength = 690 nm) emitted from the VI-910's source and the signal reflected back by the target is captured by the VI-910's CCD receiver. Coordinates (x, y, and z) of imaged objects are reconstructed through triangulation. This device stores a mesh of 640×480 3D points at each acquisition. The VI-910 is provided with three interchangeable lenses to fit a variety of scanning settings. A single acquisition captures an area between ~10 cm^2 (TELE lens) and ~0.8 m^2 (WIDE lens). The instrument maximum accuracy is achieved using the TELE lens: 0.22 mm in x, 0.16 mm in y, and 0.1 mm in z, the z axis being the optical axis of the laser scanner. For this work we used only the WIDE lens which has accuracies of 1.4 mm along x, 1.04 mm along y, and 0.4 mm along z. 3D models are created using the Konica Minolta Polygon Editing Tool by data alignment, merging, and triangulation. No successive filling or smoothing was performed.

<h1>3. Test cases

<h2>3.1. Outcrops

We selected three sites (S1, S2, and S3) exposed on a subvertical fresh outcrop in a marble quarry located on the South flank of Mt. Castellare, near San Giuliano Terme (Pisa, Italy; Figs. 2a, b, and c). From a geological perspective the area belongs to the Monti Pisani Unit, one of the main metamorphic outcrops of the Northern Apennine, that has been subjected to two main episodes of deformation, a first compressive ductile phase between the late Oligocene and the early Miocene followed by an extensional phase during the Tortonian (Carosi et al., 2004, and references therein). This poly-phase

deformation history results in a complex pattern of fractures clearly visible on the quarry surface. The three sites show various chromatic and textural characteristics representing different geological aspects, despite their collocation a few tens of meters apart.

At site S1 the liassic marble "Calcare ceroide" crops out (Rau and Tongiorgi, 1974). It is a low-grade metamorphic white, gray or whitish-yellowish marble with thin layers of muscovite (Figs. 2a and 3). The quarry cuts small cave passages unearthing physical and chemical cave deposits. At site S2 different speleothems are present (Figs. 2b and 4): (i) a thin flowstone originated as a calcite deposit from a uniform water flow and accreted roughly parallel to the surface (almost vertical in this case); (ii) a small stalactite (i.e., a subvertical concretion growing from top to bottom as a result of carbonate deposition from water drops); and (iii) cave popcorn concretions (i.e., globular calcite deposits developed in a low evaporation environment). At site S3, a small sedimentary breccia section crops out. It is an unsorted debris, grain supported, probably derived from a colluvium tongue, transported in depth by gravity through a fracture (Figs. 2c and 5).

At each site we collected a large series of pictures suitable for the multiview 3D reconstruction procedure, by using a Canon EOS 450D digital camera. The same scenes have been imaged by using the Konica Minolta VI-910 laser scanner mounted with the TELE lens. 3D models of sites have been built from both acquisition systems.

The acquired areas are approximately rectangular and cover extents between ~0.15 and ~0.3 m^2 (average linear dimensions from ~40 to ~55 cm, see Figs. 3, 4, and 5, Table 1). To explore the effectiveness of the multiview method with respect to the series of input pictures, the models of the three outcrops have been derived by processing a different number of pictures. The model for site S1 was obtained by processing four photos, producing a final cloud of ~55,000 points; the model for site S2 was obtained by processing 40 photos, resulting in a final cloud of ~200,000 points; and the model for site S3 by processing 35 photos, and a final cloud of ~450,000 points.

The point cloud density is clearly related to the number of input photos but also to their quality and to the acquisition geometry. In fact models of S2 and S3 have been reconstructed starting from a similar number of pictures (40 vs 35) but the average number of points per photo in the final point clouds is rather different (~200,000 vs ~450,000; Table 1).

<H2>3.2. 3D modeling of solids

We analyzed two solids different in shape, color, mineral composition, and geological meaning: a stalagmite and a volcanic bomb (Figs. 2e and d, respectively). A stalagmite is a speleothem growing from the floor of a cave caused by the dripping of water rich in calcium bicarbonate. A volcanic bomb is a lava projectile which by definition is larger than 65 mm in diameter and ejected by a volcano during an eruption.

The stalagmite modeled in this work was taken from the Buca di Cavorso (Jenne, Roma, Italy). It has the typical tapered shape (Figs. 1 and 6), a height of ~27 cm, and basal diameter of ~10 cm. A 3D model was reconstructed using 30 photos obtaining a cloud of ~185,000 points. Comparison with the 3D model obtained using the laser scanning is shown in Fig. 6 and tabulated in Table 2.

The volcanic bomb considered here was ejected during the 2001 eruption at Mt. Etna (Italy) from the South-East summit crater. This bomb has the typical almond shape (Fig. 7) with the maximum and minimum dimensions of ~15 and ~9 cm, respectively. By using 67 input photos we derived a cloud of ~136,000 points (Table 2).

<H1>4. Discussion

The outcrops have a simple, almost planar surface and can be reconstructed by using a small number of photos. On the contrary, the much more complex reconstruction of solids requires tens of photos. For almost flat surfaces (outcrops S1 to S3) the effective number of points per photo in the points cloud is high (5000–15,000 points/image for a 1024×638 pixel image; see Table 1); for solids, the number of points per used photo drops significantly (2000–6000 points/image) despite the simple geometry of the considered samples (Table 2).

For the error assessment, we considered as “ground” truth the 3D models obtained by using the Konica Minolta VI-910 laser scanner, owing to the low nominal error. The multiview model of S1 has been derived by using only 4 photos; nevertheless it shows a low root mean square error (RMSE), though significantly higher than the ones calculated for the models of S2 and S3. The RMSEs percentage (i.e., RMSE/average linear dimensions) is 0.68% in the model generated from four photos and 0.11% in the model generated from 35 photos, which turns out to be the most accurate. The higher error obtained in the S1 site is easily explained: S1 presents quasi-planar surfaces broken by big discontinuities and four photos are not able to reconstruct such big discontinuities

(e.g., the red area in Fig 3c). Thus a percentage RMSE of 0.68% must be considered a conservative upper limit rather than the rule since it refers to the worst possible combination of acquisition geometry and surface characteristics.

For the three outcrops, the maps of the depth differences between the models obtained with multiview 3D and the control models obtained using the laser scanning (Figs. 3, 4, and 5) show a clear pattern with positive values at the edge of the scenes and negative ones at the center. This evidence suggests the existence of a systematic error in our multiview 3D reconstruction.

We used the outcrop models (Figs. 3, 4, and 5) to quantify textural differences among the three sites. We calculated two parameters: (i) the roughness as the root mean square heights along the viewing direction, and (ii) the detrended roughness, calculated as above after the subtraction of the best fitting plane from the model. For the purpose of roughness calculations, the photo-derived triangulated surfaces are "georeferenced" with the corresponding laser-derived surfaces and then all the pairs of surfaces are converted into grids. Roughness calculations are performed on the gridded surfaces. Results show that the detrended roughness of S1 is higher than that of S2 and S3 (16.73 vs 12.56 and 9.79 mm, respectively; Table 1). Percentage errors in detrended roughness, derived by comparing photo-derived and laser-derived 3D models, are in the range 0.3–2%.

For the two solids, we used the software Metro to calculate the errors of the photo-derived models with respect to the laser-derived models. The stalagmite model has an overall RMSE of ~0.80 mm, corresponding to a percentage RMSE/average sample linear dimension of 0.22%. The volcanic bomb model has an RMSE of ~0.33 mm, corresponding to a percentage RMSE/average sample linear dimension of 0.16% (Table 2). Table 2 shows that RMS distance between the laser-derived and the photo-derived models can change significantly as the reference solid changes (i.e., the solid from which the distance is calculated according to Eq. (1)). This is due to missing portion in one of the models, for example, at the base of bomb in the photo-derived model (Fig. 7). The photo- and laser-derived models of the stalagmite are more consistent (Table 2 and Fig. 6).

Fig. 8 shows the error distributions in all the test cases. S1, S2, and S3 show an asymmetric distribution which is due to the above-described systematic error (see Figs. 3c, 4c, and 5c). Despite the apparent greater error spreading of S2 and S3, S1 has the higher RMSE, owing to the biased reconstruction of the discontinuity which cuts almost horizontally across the sampled surface in the photo-derived model (Fig. 3). For the bomb

and the stalagmite we plotted the discrepancies (always positive) between photo- and laser-derived models.

To explore the sensitivity of the method with respect to the PMVS2 settings, we iteratively rederived all our models introducing small changes in the PMVS2 option file. We found that these small changes result in negligible variations in points cloud density and model accuracy.

<h1>5. Conclusions

We assessed the performances of a multiview 3D reconstruction method for generating full 3D models of small outcrops (areas between ~0.15 and ~0.3 m^2) and decimeter-scale objects of geological interest. The complete processing is carried out by using only freely available software.

Comparisons with reference models acquired by using a laser scanner show that this method warrants percentage RMSE (RMSE/average sample linear dimension) which can attain ~0.1%. Obtained results demonstrate that the multiview 3D reconstruction technique can be effectively used to substitute much more expensive and cumbersome technologies (e.g., laser scanners or terrestrial LIDAR) in cases similar to the ones presented here. The main advantages of multiview techniques are:

- Simplicity: viable input images can be acquired without any specific competence and the final 3D reconstruction is straightforward.
- Flexibility: a multiview survey does not involve logistical efforts because it requires only the use of a digital camera (easily to bring everywhere).
- Low cost: multiview 3D reconstruction involves a consumer-grade camera, freely available software, and the survey does not require additional costs.
- Scale free: multiview methods are, in theory, not constrained in scale, as long as the acquired series of pictures fit the required specifications.
- Acquisition frequency: an acquisition can be as fast as a click: setting up several cameras in different locations, full 3D acquisition can be done at very short time steps. This can be very useful, for example, to support laboratory experiments.

On the other hand, more expensive techniques can reach higher resolutions and accuracies and/or work at much longer ranges. Also, lighting conditions affect the final result, while active acquisition systems do not have similar problems.

As a whole, photo-derived 3D reconstructions turn out to be easy, fast, reliable, and nonexpensive for 3D modeling of scenarios of geological interest. Possible future applications could include the determination of morphological changes of rapidly evolving systems (e.g., steep, unstable slopes or riverbeds) and the monitoring of in-laboratory analog experiments.

Acknowledgments

A..F benefited from the MIUR FIRB project “Piattaforma di ricerca multidisciplinare su terremoti e vulcani (AIRPLANE)” n. RBPR05B2ZJ. S.T. and I.I. benefited from the FIRB project "Sviluppo di nuove tecnologie per la protezione e difesa del territorio dai rischi naturali (FUMO)" funded by the Ministero dell'Istruzione, dell'Università e della Ricerca.

355 **References<style>**

356

357 Bellian, J.A., Kerans, C., Jennette, D.C., 2005. Digital outcrop models: Applications of
358 terrestrial scanning LIDAR technology in stratigraphic modelling. Journal of
359 Sedimentary Research 72(2), 166–176.
360 Besl, P.J., McKay, N.D., 1992. A method for registration of 3-D shapes. IEEE
361 Transactions on Pattern Analysis and Machine Intelligence 14, 239–256.
362 Buckley, S.J., Howell, J.A., Enge, H.D., Kurz, T.H., 2008. Terrestrial laser scanning in
363 geology: Data acquisition, processing and accuracy considerations, Geological
364 Society of London 165, 625–638.
365 Carosi, R., Montomoli, C., Pertusati, P.C., 2004. Late tectonic evolution of the Northern
366 Apennines, the role of contractional tectonics in the exhumation of the Tuscan unit.
367 Geodinamica Acta 17, 253–273.
368 Cignoni, P., Rocchini, C., Scopigno, R., 1998. Metro: Measuring error on simplified
369 surfaces. Computer Graphics Forum 7(2), 167–174.
370 de Matías, J., de Sanjosé, J.J., López-Nicolás, G., Sagüés, C., Guerrero, J.J., 2009.
371 Photogrammetric methodology for the production of geomorphologic maps:
372 Application to the Veleta Rock Glacier (Sierra Nevada, Granada, Spain). Remote
373 Sensing 1, 829–841.
374 Dowling, T.I., Read, A.M., Gallant, J.C., 2009. Very high resolution DEM acquisition at
375 low cost using a digital camera and free software. In: Proceedings of the 18th World
376 IMACS / MODSIM Congress, Cairns, Australia.
377 Furukawa, Y., Ponce, J., 2007. Accurate, dense, and robust multi-view stereopsis. In:
378 Proceedings, IEEE Conference on Computer Vision and Pattern Recognition CVPR
379 2007, pp. 1–8.
380 Furukawa, Y., Ponce, J., 2009. Accurate camera calibration from multi-view stereo and
381 bundle adjustment. International Journal of Computer Vision 84, 257–268.
382 Grzeszczuk, R., Košecka, J., Vedantham, R., Hile, H., 2009. Creating compact
383 architectural models by geo-registering image collections. In: Proceedings of the 2009
384 IEEE International Workshop on 3D Digital Imaging and Modelling (3DIM 2009),
385 Kyoto, Japan.
386 Hartley, R.I., Zisserman, A., 2004. Multiple View Geometry in Computer Vision,
387 Cambridge University Press.

James, M.R., Pinkerton, H., Robson, S., 2007. Image-based measurement of flux variation in distal regions of active lava flows. Q03006. doi:10.1029/2006GC001448.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., Fulk, D., 2000. The Digital Michelangelo Project: 3D scanning of large statues. Computer Graphics (SIGGRAPH 2000 Proceedings).

Lourakis, M., Argyros, A., 2008. SBA: A generic sparse bundle adjustment C/C++ package based on the Levenberg-Marquardt algorithm. http://www.ics.forth.gr/lourakis/sba.

Lowe, G.D., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110.

McCaffrey, K.J.W., Jones, R.R., Holdsworth, R.E., Wilson, R.W., Clegg, P., Imber, J., Holliman, N., Trinks, I., 2005. Unlocking the spatial dimension: Digital technologies and the future of geoscience fieldwork. Journal of the Geological Society 162(6), 927–938.

Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to Modern Photogrammetry, John Wiley & Sons, Inc., New York.

Nister, D., 2004. Automatic passive recovery of 3D from images and video. In: Proceeding of the 2nd IEEE International Symposium on 3D Data Processing, Visualization and Transmission, pp. 438–445.

Pringle, J.K., Howell, J.A., Hodgetts, D., Westerman, A.R., Hodgson, D.M., 2006. Virtual outcrop models of petroleum analogues: A review of the current state-of-the-art. First Break 24(3), 33–42.

Rau, A., Tongiorgi, M., 1974. Geologia dei Monti Pisani a sud-est della valle del Guappero. Memorie Società Geologica Italiana 8, 227–408.

Snavely, N., Seitz, S.M., Zeliski, R.S., 2006. Photo tourism: Exploring image collections in 3D. In: SIGGRAPH Conference Proceedings, New York, NY, USA, ACM Press, pp. 835–846.

Snavely, N., Seitz, D., Szeliski, R., 2007. Modeling the world from internet photo collections. International Journal of Computer Vision 80, 189–210.

Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 2000. Bundle adjustment—A modern synthesis. In: Triggs, W., Zisserman, A., Szeliski, R. (Eds), Vision Algorithms: Theory and Practice, LNCS. Springer–Verlag, pp. 298–375.

422

423 **Figure captions**

424

425 **Fig. 1.** Camera positions and orientations used in the acquisition of the stalagmite. (a) Top
426 view; (b) lateral view.

427

428 **Fig. 2.** Surfaces used as test cases for the generation of 3D models: (a,b,c) outcrops
429 exposed on a subvertical fresh wall in a marble quarry located on the South flank of Mt.
430 Castellare, near San Giuliano Terme (Pisa, Italy; surfaces S1, S2, and S3 of Figs. 3, 4, and
431 5, respectively, are outlined by red dashed lines); (d) stalagmite from the Buca di Cavorso
432 (Jenne, Roma, Italy; Fig. 6); (e) volcanic bomb ejected during the 2001 eruption at Mt.
433 Etna (Italy; Fig. 7).

434

435 **Fig. 3.** Digital model of San Giuliano marble outcrop (site S1). (a) 3D point cloud from
436 the multiview reconstruction, displayed by RGB color information; (b) slope model of
437 laser-derived data; (c) difference map between multiview reconstruction and laser-derived
438 model; (d) slope model of the multiview reconstruction.

439

440 **Fig. 4.** Digital model of calcareous concretion outcrop (site S2). (a) 3D point cloud from
441 the multiview reconstruction, displayed by RGB color information; (b) slope model of
442 laser-derived data; (c) difference map between multiview reconstruction and laser-derived
443 model; (d) slope model of the multiview reconstruction.

444

445 **Fig. 5.** Digital model of small breccias outcrop (site S3). (a) 3D point cloud from the
446 multiview reconstruction, displayed by RGB color information; (b) slope model of laser-
447 derived data; (c) difference map between multiview reconstruction and laser-derived
448 model; (d) slope model of the multiview reconstruction.

449

450 **Fig. 6.** 3D model of a stalagmite: (a) model derived from the multiview reconstruction,
451 displayed by RGB color information; (b) shaded image of laser-derived model; (c) 3D
452 difference map between multiview-derived and laser-derived surfaces.

453

454 **Fig. 7.** 3D model of a volcanic bomb: (a,d) model derived from the multiview
455 reconstruction, displayed by RGB color information; (b,e) shaded image of laser-derived
456 model; (c,f) 3D difference map between multiview-derived and laser-derived surfaces.
457
458 **Fig. 8.** Error distributions of multiview-derived models of the outcrops (S1, S2, and S3),
459 the stalagmite and the volcanic bomb considered in this work. Errors are evaluated as
460 differences with laser-derived models. For the stalagmite and the volcanic bomb, errors
461 are evaluated as distances between the multiview-derived and the laser-derived surfaces
462 (Eq. (1)).
463
464

<Comp: please remove internal rules from tables>

Table 1

Characteristics of sampled outcrops, laser-derived and multiview-reconstructed models.

	Parameter	S1	S2	S3
Outcrop extension	Area (m^2)	0.306	0.169	0.148
	X extent (mm)	642	471	471
	Y extent (mm)	476	359	315
	Average XY scale[a] (mm)	553	411	385
Laser model	N. pts.	269390	226172	189744
	Average mesh step (mm)	1.07	0.86	0.88
	Roughness (mm)	25.90	31.34	16.92
	Detrended roughness (mm)	16.73	12.56	9.79
Photo model	N. photo	4	40	35
	N. pts.	55265	205252	450075
	Average N. pts./N. photo	13816	5131	12859
	Average mesh step (mm)	2.35	0.91	0.57
	Roughness (mm)	25.36	28.89	16.90
	Detrended roughness (mm)	16.68	12.24	9.67
	RMSE[b] (mm)	3.76	1.09	0.41
	Percentage error[c] (%)	0.68	0.27	0.11

[a] Calculated as the square root of area.

[b] Root mean square error between the laser-derived 3D model and the multiview 3D reconstruction.

[c] Calculated as the ratio between the RMSE and the average XY scale.

Table 2

Characteristics of laser-derived and multiview-derived models and distance between the two surfaces.

Parameter	Stalagmite		Volcanic bomb	
	Laser	Photo	Laser	Photo
N. vertices	137848	185628	82413	136519
N. faces	269677	320887	156414	236039
Area (mm^2)	84969	83392	39683	31708
Bounding box diag. D (mm)	368	423	214	204
Max distance (mm)	12.8	14.2	17.6	4.5
Mean distance (mm)	0.54	0.55	0.75	0.23
RMS distance E (mm)	0.81	0.92	2.14	0.33
E/D (%)	0.22	0.22	1.00	0.16

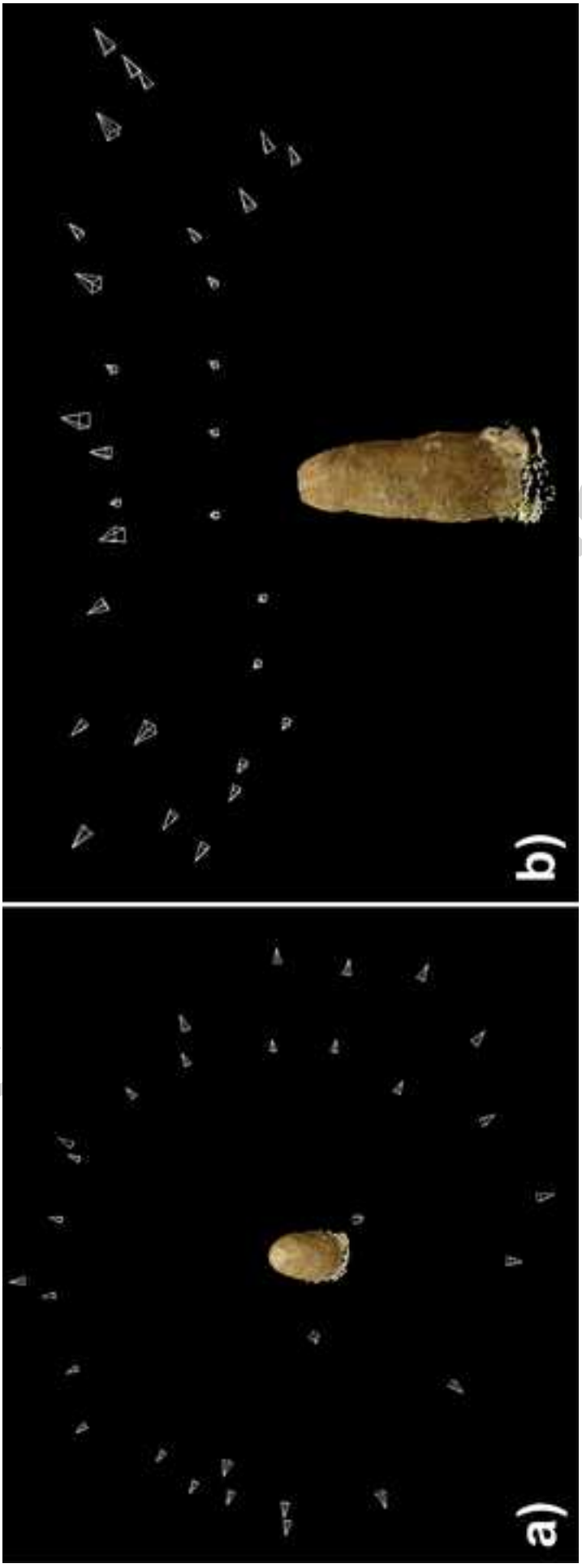


Figure 1

Figure 2

Figure 3

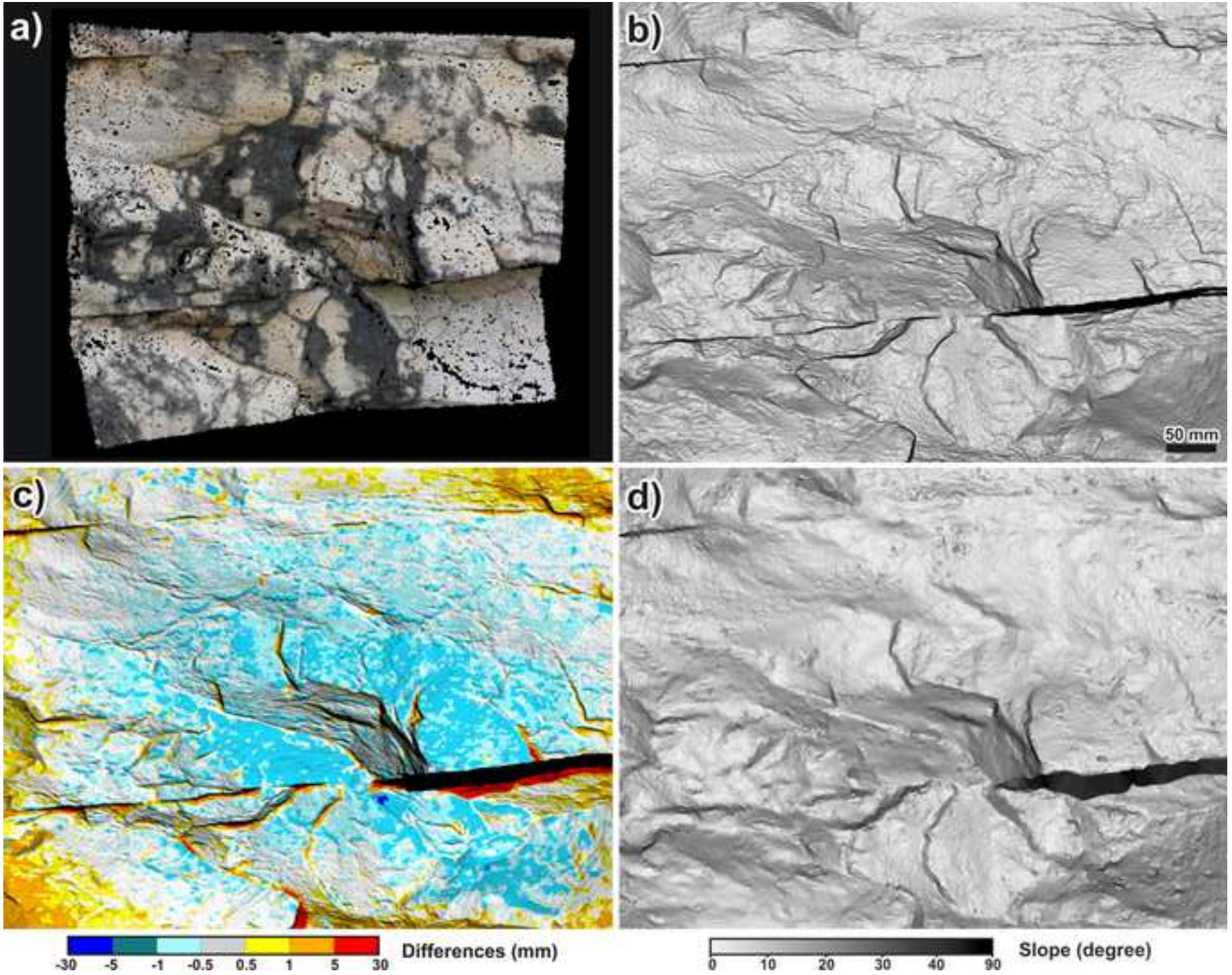


Figure 4

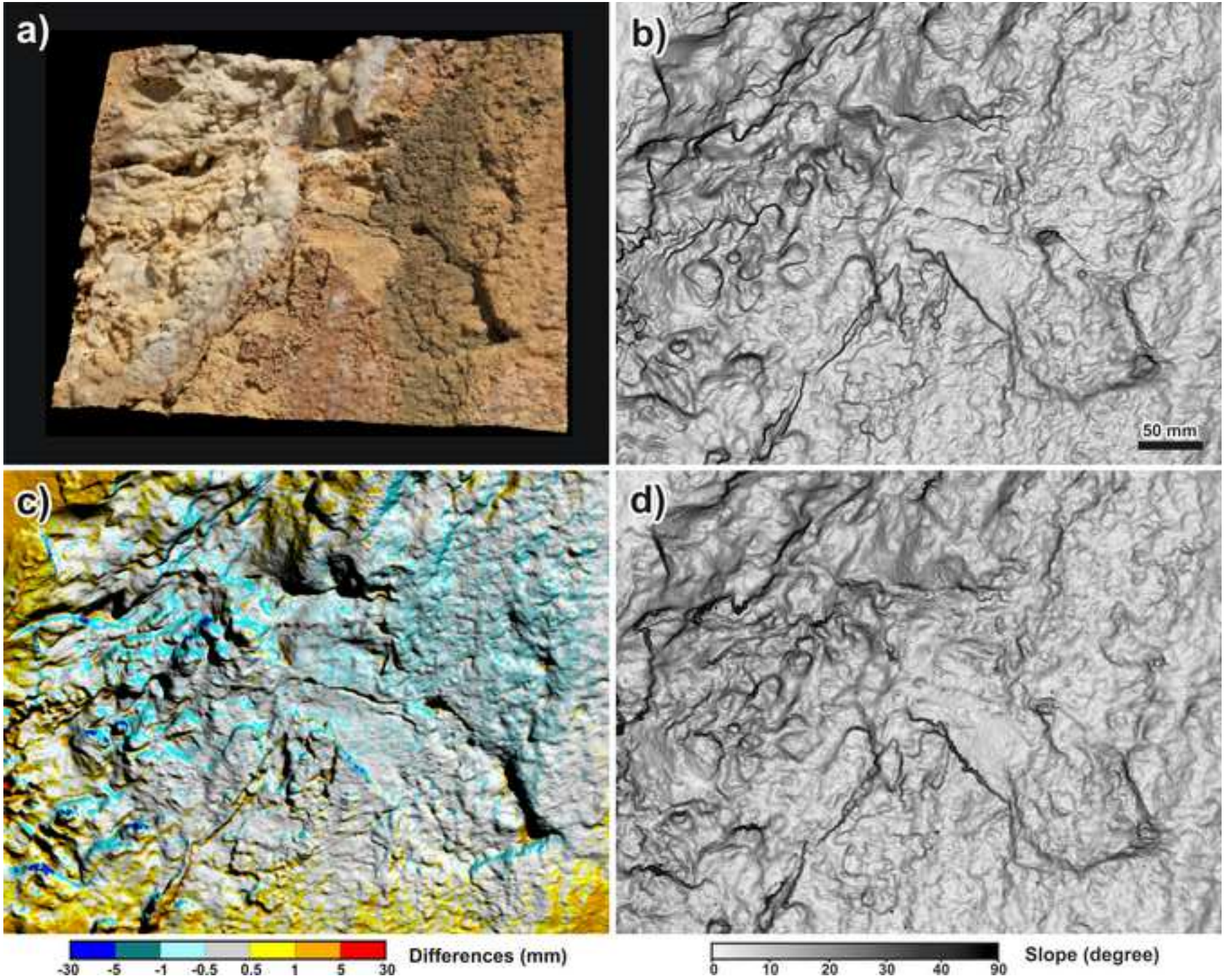


Figure 5

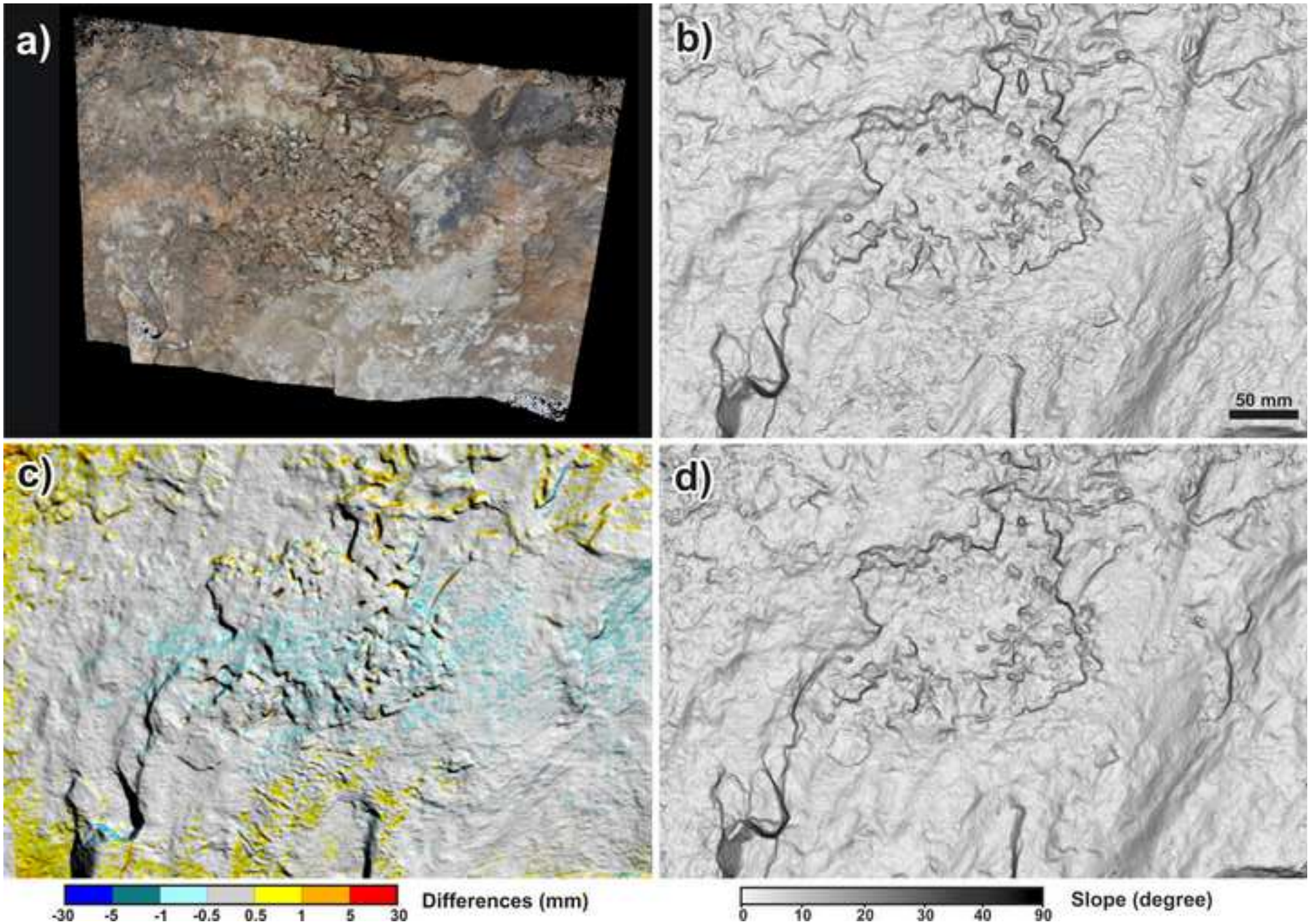
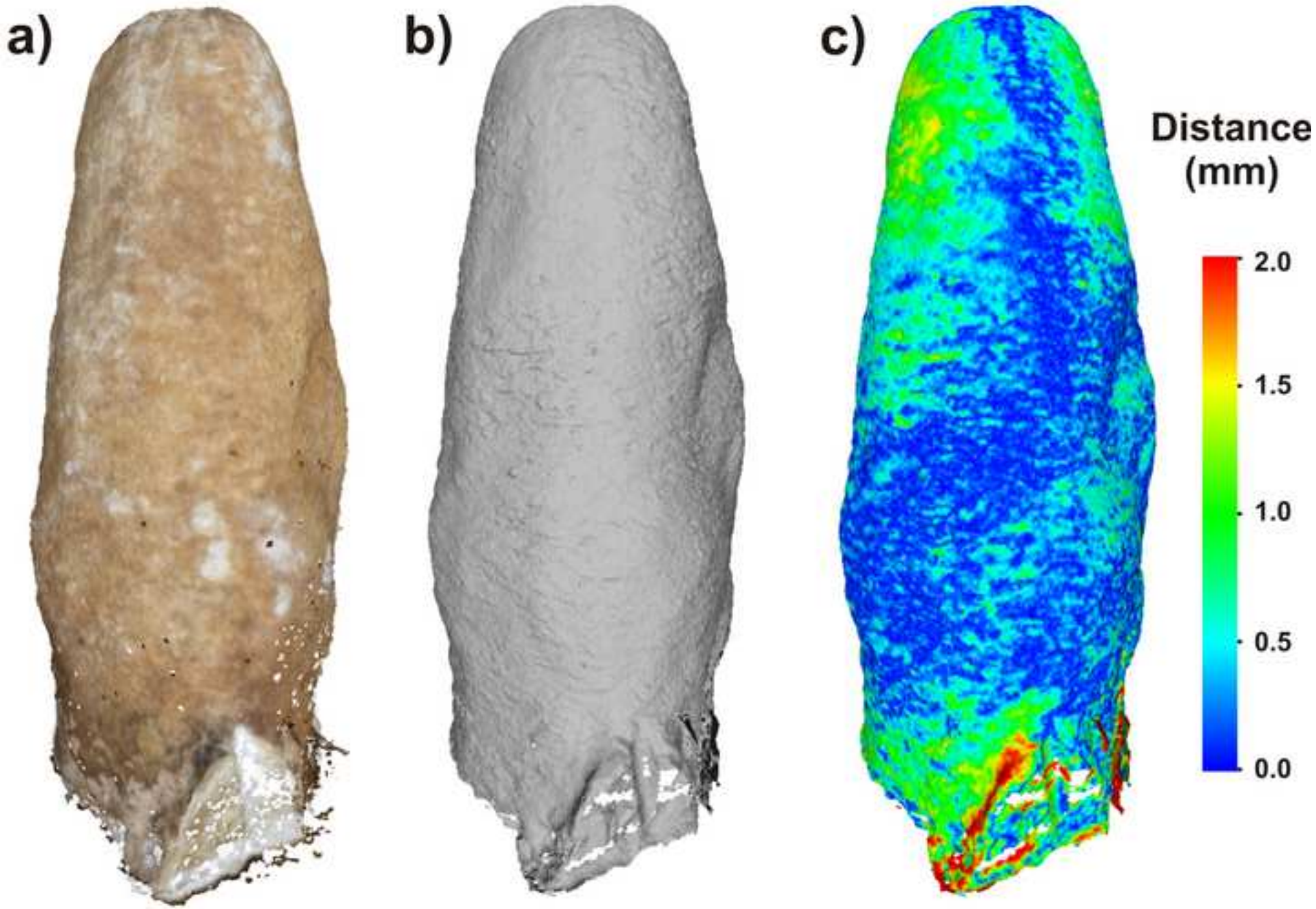


Figure 6



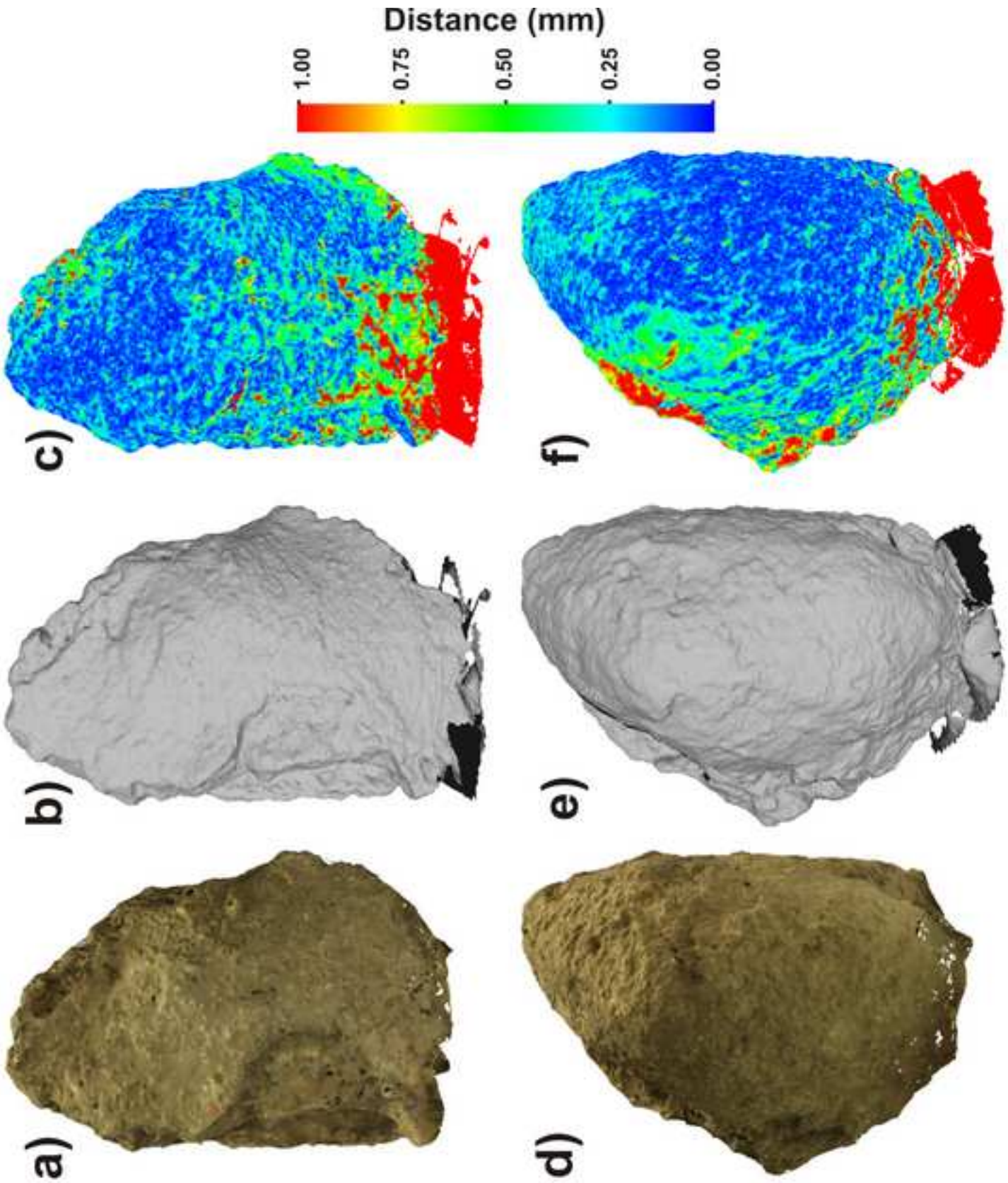


Figure 7

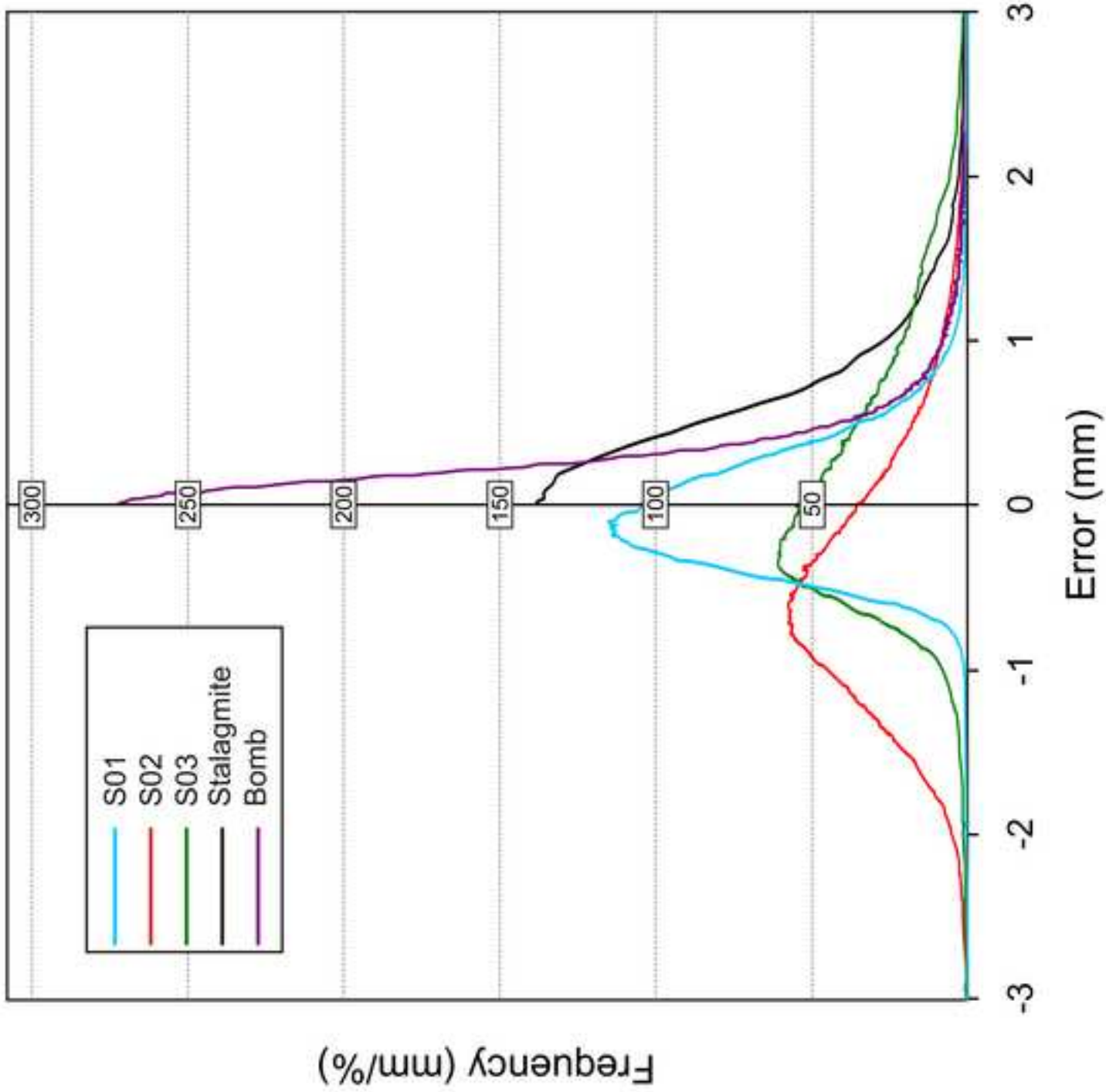

Frequency (mm/%)
300
250
200
150
100
50
S01
S02
S03
Stalagmite
Bomb
-3
-2
-1
0
1
2
3
Error (mm)

Figure 8