**ORIGINAL PAPER** 



# Methodologies for assessing the quality of 3D models obtained using close-range photogrammetry

Andrea di Filippo<sup>1</sup> · Sara Antinozzi<sup>1</sup> · Nicola Cappetti<sup>2</sup> · Francesco Villecco<sup>2</sup>

Received: 30 July 2022 / Accepted: 26 June 2023 © The Author(s) 2023

#### Abstract

Although reality-based models are widely used to describe the geometric surfaces of an entity in a digital space, a systematic and universally recognised treatment of issues such as accuracy is lacking. The topic is certainly complex as this analysis should involve not only shape approximation but also other attributes (e.g., colour). Wanting to limit ourselves to geometry alone, this work proposes solutions for assessing the quality of photogrammetric models, differentiating them according to possible scenarios: sometimes, homologous models obtained using different techniques and technologies are available. In these cases, a comparison between digital reconstructions can serve to effectively quantify accuracy; more often, no terms of comparison are available, and one is forced to derive indicators from the same photogrammetric process to describe quality. We propose for this scenario a statistical analysis on the covariance matrix of the estimated coordinates for the tie points. The main goal is to provide a range of possible approaches to the conscious management of survey data.

Keywords Tolerance intervals · Coordinate covariance matrix · Hausdorff distance · Accuracy assessment

# 1 Introduction

The use of reality-based models, derived from techniques such as laser scanning or photogrammetry, is widely used in fields ranging from heritage documentation to reverse engineering on mechanical components [1, 2].

The need for realism and detail has produced, over time, a continuous increase in the complexity and size of virtual descriptions, challenging the storage, transmission, and processing capacity of hardware.

Given these assumptions, it seems clear that the properties of a digital outputs must be calibrated, both in terms of richness of detail and accuracy, in relation to the specific goals. Despite intensive use of models, whether in the form of point clouds or polygonal meshes [3], there is no agreement on the most appropriate method for defining their quality, and a common criterion for formalizing error is lacking.

Francesco Villecco fvillecco@unisa.it This is because computing processes, and in particular algorithms for handling raw data, can differ widely [4].

The main goal of this paper is to provide effective solutions for quantifying the reliability of a photogrammetric description in relation to geometric attributes. We limit ourselves to close-range applications [5], investigating two opposite scenarios:

- in the former case we have, in addition to the photogrammetric model, a homologous digital object obtained by more reliable and accurate processes or tools. It is then possible to use the latter as a reference and perform a comparison to analyse the distribution of distances between the two entities. The steps to be followed must obviously be detailed according to the features of the investigated items;
- in the second one, there are no supporting digital descriptions and therefore quality indicators must be derived directly from the photogrammetric process, which requires strict control of all parameters governing it.

The proposed approaches provide maximum flexibility and are applicable to both point clouds and meshes, regardless of the algorithms used to produce them. In the following

<sup>&</sup>lt;sup>1</sup> Department of Civil Engineering, University of Salerno, Fisciano, Italy

<sup>&</sup>lt;sup>2</sup> Department of Industrial Engineering, University of Salerno, Fisciano, Italy

paragraphs they will be duly elaborated and tested with reference to a vase of ceramic material, surveyed by both photogrammetric technique and structured light scanner. The two scenarios outlined before will be treated appropriately, highlighting the possible critical issues, and differentiating the procedures according to the specifics of the case.

# 2 Background

#### 2.1 Comparing homologous models

It often happens that, in processes of digitizing an object, it is detected and returned (totally or partially) using multiple techniques. There can be many reasons for this redundancy of information, ranging from data integration (at any level) to the production of multi-resolution outputs [6].

Regardless of the purpose, it is possible to exploit this content to quantify the accuracy of geometric attributes. Among the homologous models, there will be one that is more reliable both in terms of how the data is acquired and how it is processed, depending on the employed technique. This description can be used as a reference, calculating the distances of the other models from it and studying the discrepancy distributions with appropriate statistical tools.

However, the intuitiveness of the process hides a complexity that depends largely on the way the different products are obtained, which is why we do not have a one-size-fitsall solution in the literature to perform a comparison. For clarity, it is appropriate to report the questions that underlie many disagreements on the procedure to be followed:

- How should we choose the reference entity? (i)
- How should models be registered? (ii)
- What characteristics should digital descriptions possess? (iii)
- What is the most appropriate algorithm for quantifying distances? (iv)

Starting with the first question (i), the most reliable model should be used as a reference, but its determination is not unambiguous and depends on many factors. The foremost is the resolution, i.e. the level of recognition of the smallest variation in magnitude on the surface of the digitized object. This is followed by the accuracy of the campaign, related to acquisition and processing techniques. For example, the proposed applications include a comparison between a model obtained from a single scan with a structured light instrument and a resolution of 0.2 mm, used as a reference, and a photogrammetric one with a Ground Sample Distance (GSD) of 0.4 mm/pixel. In general, active optical sensors are more reliable from a purely geometric perspective. This

depends not only on the inherent features of the instrumentation but also on the fact that the operator, in both the detection and handling phases of the raw data, has to control a very limited number of factors, internal and environmental, compared with a process involving passive sensors. Furthermore, photogrammetric acquisition is clearly separated from the data management step, the variables of which can greatly affect the final product according to subjective choices by technicians. A model generated with manufacturer-declared specifications with almost no influence by personal preferences would therefore provide a reliable basis for comparison.

For the definition of a common reference system (ii), we prefer a punctual approach using the coordinates of appropriate targets positioned in the surveyed scene. Solutions that involve entire digital objects, such as ICP-derived algorithms [7], are certainly more robust but, as they do not rely on homological relationships between points, they attempt to reduce distances by searching for the configuration that ensures the best overlap, running the risk of localized anomalies not being duly revealed.

About the features of the compared models (iii), it is preferable that they have a similar resolution or surface density, or at least a higher one for the reference entity, since almost all algorithms available to quantify discrepancies are based on the calculation of normals or local modelling applied to the latter, which is fundamental to the success of the analysis.

Linked to the previous argument is the choice of the algorithm (iv). If you have a reference in the form of a polygonal mesh, you can use the Cloud-to-Model (C2M) distance [8, 9]. This approach is the most common technique in inspection. Surface change is calculated by the distance between a point cloud and a reference 3D mesh or theoretical model. It works well on flat surfaces, as a mesh corresponding to the average reference point cloud position can be constructed [10].

This approach is not always convenient. Creating a surface mesh is complex for point clouds with significant roughness at all scales or missing data due to occlusion. The process of creating a surface could smooth out some details that may be important to assess local roughness properties. In other cases, the interpolation over missing data introduces uncertainties that are difficult to quantify, requiring time-consuming manual inspection.

Therefore, it is preferable to use the Multiscale Model-to-Model Cloud Comparison (M3C2) algorithm, which, with its parameters, allows better control of sources of uncertainty. This algorithm combines three crucial elements: it performs directly on point clouds without meshing; it calculates the local distance between two point clouds along the surface normal direction (i.e. considering 3D variations in surface orientation); it estimates for each distance measurement a confidence interval that depends on the roughness of the point cloud and the registration error [11].

The M3C2 algorithm calculates a local average cloudto-cloud distance for a point in the reference cloud, termed the core point, through the use of a search cylinder projected along a locally oriented normal vector [11]. Then, the distance is assigned as an attribute of the core point. The entire reference cloud can be defined as core points, or a subsampled set of the reference cloud. The original resolution of both point clouds is used in the M3C2 computations, regardless of whether the data is subsampled in the process. The core point's normal vector is estimated from its surrounding neighbourhood, which should be of a scale such that it captures the surface geometry without being sensitive to local surface roughness [11]. Points encompassed by the search cylinder are used to compute the average position of the compared clouds. The distance between the average positions (along the normal vector) is the M3C2 distance. The projection diameter size and the maximum search length are chosen based on the application, point spacing, and surface complexity [12].

The analysis of distance distributions requires additional statistical tools. Very effective are the tolerance intervals [13], which allow the nature of the distributions and the size of the samples to be properly considered. The above considerations provide good coverage of the possible scenarios that arise during operations.

Before proceeding further, it is worth mentioning that a more reliable reference object cannot always be found. In this case, it is preferable to derive, from the distance distributions, the mutual surpluses, and then calculate the Hausdorff distance [14, 15] between the two entities. However, this approach requires prior filtering to remove outliers, which can be performed, for example, with a box plot [16, 17].

# 2.2 Direct accuracy assessment of photogrammetric models

The evaluation of a photogrammetric model accuracy is more difficult without a reference object for a comparison, since quality indicators must be derived directly from the data processing. In the structure estimation and optimization steps, which include the internal and external (relative and absolute) orientation of the frames, it is essential to know the coordinates or dimensions of certain elements arranged in the object space (the scene to be digitized).

In applications of very close-range photogrammetry, it is common practice to add to the scene punctual objects with known coordinates (targets). Although simply natural image features or textures can be used for this stage, the employment of artificial targets is encouraged by faster computation times and higher accuracy in the recognition. Their purpose, therefore, would be to define a local coordinate system, to scale of the model and highlight real matches to improve the photo alignment procedure.

However, targets – in the form of Ground Control Points (GCPs) or Check Points (CPs) – are often used as the basis for analysing the accuracy of the model. Despite a good distribution across the scene, this is not a robust strategy, as their number cannot be compared with the multitude of points from which the final product is composed.

There are then additional critical issues. Let's start with CPs, whose coordinates are not directly used to optimize the model structure, but only to perform an a posteriori check on the output. We could think of combining the error associated to the CPs - expressed through the discrepancy between input target coordinates and those estimated by the photogrammetric process - and the error that characterizes the definition of target coordinates in object space (e.g., the tolerance on the production of a scale bar). Unfortunately, there is correlation among the two sources of uncertainty, and we cannot apply a simple propagation law. One would then have to go into the nature of this correlation to solve the problem rigorously, which is far from straightforward. These considerations can be extended to GCPs, with the aggravation that their 3D coordinates are used to solve the Bundle Block Adjustment. It is therefore expected that the error associated with them will be less than that of any other constituent point in the model, since the geometric structure is built and optimized precisely around their coordinates.

Robust approaches should consider more points. Certainly, one solution would be to consider the Tie Points (TPs) obtained from the orientation phase. However, such an approach requires a very strict control of the parameters governing the photogrammetric process. First, the accuracy of target coordinates in object space (how the 3D coordinates are measured or estimated) and in image space (how the targets are identified in the photos). Then follows the accuracy of TPs also in image space, which is mainly related to the quality of the acquisition campaign. Such an approach will be detailed in the next section.

#### 3 Materials and methods

As anticipated, our analysis is conducted on a close-range photogrammetric process applied to a ceramic material vase. The average GSD of the dataset is 0.4 mm and the project consists of 96 images with  $6000 \times 4000$  pixels, acquired with entry-level Nikon D3300 SLR camera. The lens is an

18–55 zoom set at 30 mm throughout the campaign, adjusting the focus beforehand to preserve the acquisition main distance.

We can then proceed with the orientation. Agisoft Metashape, the software used in our applications, detects points in the source photos which are stable under view-point and lighting variations and generates a descriptor for each point based on its local neighbourhood. These entities are used later to identify correspondences across the photos. This is like the well-known Scale-Invariant Feature Transform (SIFT) approach, but uses different algorithms for a higher alignment quality (feature detection). The software then applies a greedy procedure to find approximate camera locations (feature matching) and refines them later using a Self-Calibration Bundle Block Adjustment (structure estimation) [18, 19]. The latter solves the problem of internal and relative external orientation at the same time.

The next step is to import outward references to optimize the frame installation and to solve the problem of absolute external orientation. We use bars whose extremes consist of uncoded targets, which are also easily identifiable as reference elements, called "Markers". Metashape can locate them automatically, simply by choosing the type of artificial object placed in the scene (in this case uncoded crossshaped targets). After recognition, we proceed with a visual check and eventual optimization. At this stage, we select the appropriate local reference system.

We can then optimize the alignment (structure optimization). The goal is to obtain only high-quality tie points and repeatedly improve the camera model. This is the most subjective section of the workflow, and testing how many points can be removed at each stage may be necessary to create a successful product.

The error reduction phase relies on robust tie point and marker accuracy estimates, referred to the image coordinates quality (and, of course, on the accuracy of the reference elements in the object space). The proportion between these two parameters distributes the weight given to markers and tie points in the whole process [20]. Correct reference settings inputs prevent misleading statistics while an incorrect estimate of them generate a not representative error models, with lens coefficients that are very sensitive to these parameters.

Parallel to photogrammetric processing, a homologous model of the vessel is made with a structured light scanner capable of 0.2 mm resolution, which is essential for the development of the section for comparison.

#### 3.1 Comparison-based workflow

For tests on the proposed methodology (Fig. 1), we use two homologous models. The reference is a polygonal mesh obtained from a single scan with a structured light in-strument. The model compared is instead a dense point cloud obtained through the photogrammetric process. The vertices of the bar scales (9 points) distributed in the digitized scene are used to register the models, avoiding approaches derived from the ICP for the reasons explained in the Sect. 2.1.

Since the scanner directly returns a mesh, in this case we can employ the C2M algorithm to quantify the distances between the two entities. The same procedure is applicable if the object being compared is also a mesh. In this case, only the vertices are considered. If the vertices of this mesh are too scattered, it may be convenient to subsample the model with a surface density approaching that of the reference.

In all other cases, it is preferable to use the M3C2, which is much more refined and can isolate distances between objects more effectively, purifying them of the registrationdependent component. In the case of distances between digital entities, it is worth mentioning that we work with signed quantities.

To summarise the distance distributions, we use tolerance intervals, according to a procedure detailed in the Sect. 3.2. A similar approach can be followed even in the absence of a superordinate reference, considering the two entities as equally reliable and deriving the distributions of each other's distances. From these, we can derive tolerance intervals or, alternatively, the Hausdorff distance as a punctual indicator. The latter provides an interesting measure of the proximity between two digital objects, indicating the maximum distance of each point of the first one from the other: therefore, it can be more effective than the minimum distance, which completely neglects their spatial configuration, to which, instead, the Hausdorff distance is sensitive. In this case, it is necessary to perform a prior treatment of the distance distributions to eliminate any outliers, such as by constructing a box-plot diagram.

#### 3.2 Direct assessment procedure

The search for a robust approach to direct accuracy estimation led us to an observation: downstream of Bundle Block Adjustment with self-calibration, it is possible to derive for each Tie Point a covariance matrix, representative of the uncertainties associated with the estimation of its coordinates in object space.

From this matrix, it is possible to derive an ellipsoid, i.e. a region of space to which corresponds a certain probability of containing the theoretical mean value of the estimated coordinates. This is possible by performing an orthogonal diagonalisation (the spectral theorem is valid) where the eigenvalues represent the lengths of the semi-axes of the mentioned ellipsoid and the eigenvectors their directions.

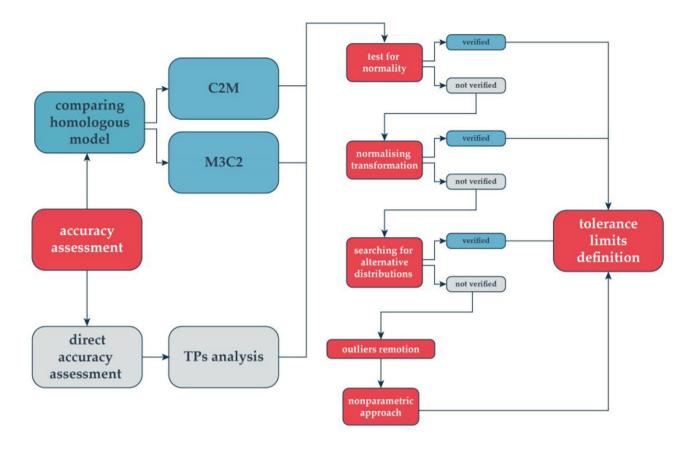


Fig. 1 Diagram summarizing the two proposed methodological approaches

Following this path presupposes that there is strict control over the data processed.

On the assumption that the control of the input data accuracy (in the object space and image space) is fundamental for a rigorous photogrammetric process, we have prepared a special Python script, with which we are able to export, after the orientation optimization phase, the covariance matrices associated with the coordinates estimated for the TPs in the object space.

Many commercial software implements a similar tool that returns an uncertainty vector obtained by composing the semi-axis lengths of the error ellipsoid with k=1. This solution is not very cautious since the probability that the theoretical mean value of the Tie Point coordinates falls in this region is 19.95%.

Instead, we consider the ellipsoid with k=3, with probability greater than 95%, and study the length distribution of its major semi-axis to derive an accuracy indicator. The statistical tools used for this analysis are tolerance intervals. They allow us estimating, from a sample, the extremes that contain a certain percentage of a population with a specific level of confidence. In this case the quantities involved are defined as positive, so we construct one-sided gaps.

The procedure can be outlined in the following steps:

- test for normality;
- search for normalising transformation (when the distribution is not normal);
- alternative distributions (when transformation approach fails);
- if all approaches fail, we calculate nonparametric tolerance limits (removal of outliers).

In general, we construct tolerance intervals with 95% confidence and 95% population. Only one parameter between the population percentage and confidence value can be defined in the case of non-parametric tolerance limit computations, with the other being determined later in the process. In that case, preventive treatment of the distributions may be necessary to eliminate possible outliers, e.g., by constructing a box-plot diagram. The approach therefore seems reasonable, considering that the TPs will constitute only a part of the final photogrammetric cloud. To be fair, the dense image matching phase and its algorithms should also be involved, but this would become too complicated. We will therefore limit ourselves to using the results of the Structure from Motion step here.

# 4 Results

# 4.1 Comparison

After registering the mesh model obtained from the scanner (reference) and the photogrammetric dense cloud (comparison) using the extremes of the bar scales, we used the C2M algorithm to obtain a distribution of distances between the two digital objects (Fig. 2). This distribution actually has a normal pattern, as verified by a special test. Considering then that the algorithm returns oriented distances, we constructed a two-sided tolerance interval with a confidence level of 95% and a population percentage of 95%, obtaining the value of -0.78 mm as the lower limit and +0.35 mm as the upper limit. Wanting to reason in absolute terms, we can take the value 0.78 mm as a synthetic indicator.

## 4.2 Direct assessment

Figure 3 shows step by step the results of the direct assessment procedure. The test for normality is not satisfied for the distribution of major semi-axes relative to ellipsoids with k=3, as can be quickly seen from the Q-Q Plot and as verified rigorously through the Shapiro-Wilk test; this leads

to the attempt to normalise the distribution. Unfortunately, a suitable power transformation could not be found; the same applies to fitting a different distribution. Non-parametric tolerance limits approach is finally performed, after removing the outliers from the distribution by constructing a box-plot diagram. Precisely this latter approach is applied to the case study fixing a confidence level of 95%, and step 4 of Fig. 3 shows the results. Since the semi-axis length is a positive definite quantity, we employ a one-sided interval, obtaining an upper tolerance limit of 2.22 mm, which can be used as an indicator of the accuracy of the entire photogrammetric process.

# **5** Discussion and conclusions

The issue of traceability of data from photogrammetric surveying is as timely as ever. Given the numerous parameters that govern the processing process, there is still a lack of systematic treatment of the topic and no agreement on how to quantify error.

The main objective of this study is to provide a response to this need, capable of ensuring maximum flexibility and adaptability to the needs of the specific case study.

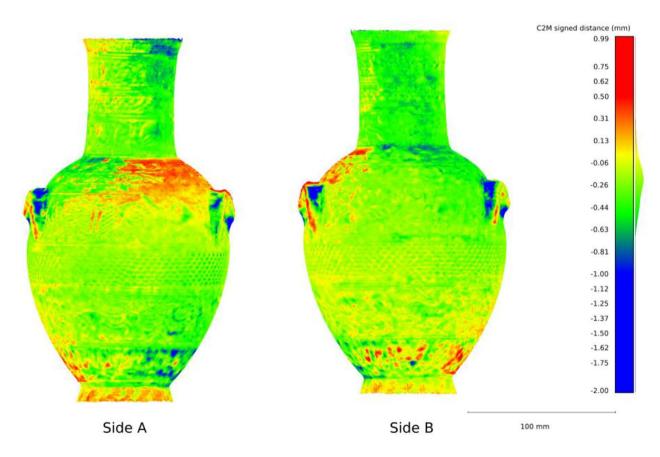
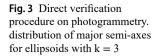
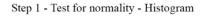
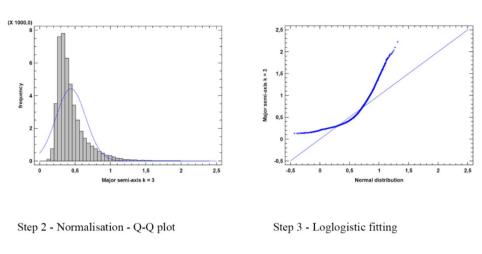


Fig. 2 C2M distance between photogrammetric (compared) and scanner model (reference)

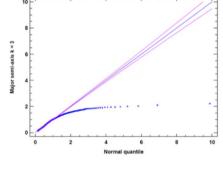


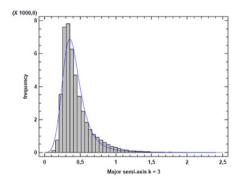


Step 1 - Test for normality - Q-Q plot



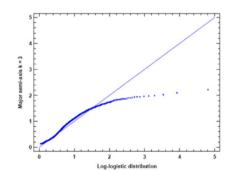


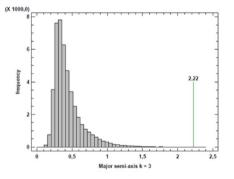




Step 3 - Loglogistic fitting - Q-Q plot

Step 4 - Nonparametric upper limit





The proposal is differentiated and follows two possible scenarios. On the one hand, the possibility of having a more reliable homologous model, obtained, for example, by laser scanning techniques. In this case, the assessment on accuracy is based on a comparison. In procedures of this kind, it is of paramount importance to consider aspects such as the topology of the models, their nature, how they are aligned, and how they are compared, to obtain robust results that can be compared effectively. Tolerance intervals are the statistical tools primarily used to describe distributions of distances because they are not bound by assumptions about their type and are sensitive to the sample size analysed. Of course, different approaches involving metrics such as Hausdorff distance can also be used, a sign of the great flexibility of the procedure. The second scenario excludes the presence of superordinate models for accuracy and focuses directly on the photogrammetric process. By appropriately calibrating the weights of all input data, we derive the covariance matrices associated with the estimated coordinates for the TPs to construct appropriate error ellipsoids, studying the distributions of their major semi-axes and extending the results to the entire model through the tolerance intervals.

The two procedures described can be used regardless of the nature of the outputs, whether point clouds or meshes. Moreover, they are not limited to the photogrammetric technique but can also be extended to laser scans or otherwise exploiting active optical sensors. Future developments will focus precisely on the applicability of the methodology to scenarios not yet investigated with the case study, proposing appropriate implementations for adaptation to the specificity of the case.

Funding Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons. org/licenses/by/4.0/.

## References

- Guidi, G., Russo, M., Angheleddu, D.: 3D survey and virtual reconstruction of archeological sites. Digit. Appl. Archaeol. Cult. Herit. 1(2), 55–69 (2014)
- Grilli, E., Remondino, F.: Classification of 3D Digital Heritage. Remote Sens., 11(7). (2019)
- Denning, J.D., Kerr, W.B., Pellacini, F.: MeshFlow: interactive visualization of mesh construction sequences, in ACM SIG-GRAPH 2011 papers. 2011, Association for Computing Machinery: Vancouver, British Columbia, Canada. p. Article 66
- Morena, S., Barba, S., Álvaro-Tordesillas, A.: SHINING 3D EINSCAN-PRO, APPLICATION AND VALIDATION IN THE FIELD OF CULTURAL HERITAGE, FROM THE CHILLIDA-LEKU MUSEUM TO THE ARCHAEOLOGICAL MUSEUM OF SARNO. Int. Arch. Photogramm Remote Sens. Spatial Inf. Sci. XLII-2/W18, 135–142 (2019)

- 5. Luhmann, T., et al.: Close-Range Photogrammetry and 3D Imaging. De Gruyter (2013)
- Ramos, M.M., F.J.T.I.A.o, P., Remondino, R., Sensing, Sciences, S.I.: Data fusion in cultural heritage-A review 40(5): p. 359. (2015)
- Chetverikov, D., et al.: The Trimmed Iterative Closest Point algorithm. in 2002 International Conference on Pattern Recognition. (2002)
- Guerin, A., et al.: Brief communication "Estimating rockfall frequency in a mountain limestone cliff using terrestrial laser scanner" Nat. Hazards Earth Syst. Sci. Discuss., 2014: p. 123–135. (2014)
- Kromer, R.A., et al.: Identifying rock slope failure precursors using LiDAR for transportation corridor hazard management. Eng. Geol. 195, 93–103 (2015)
- D'Amato, J., et al.: Influence of meteorological factors on rockfall occurrence in a middle mountain limestone cliff. Nat. Hazards Earth Syst. Sci. 16(3), 719–735 (2016)
- Lague, D., Brodu, N., Leroux, J.: Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). ISPRS J. Photogrammetry Remote Sens. 82, 10–26 (2013)
- Nourbakhshbeidokhti, S., et al.: A workflow to Estimate Topographic and volumetric changes and errors in Channel Sedimentation after Disturbance. Remote Sens. 11 (2019). https://doi. org/10.3390/rs11050586
- 13. Natrella, M.G.: Experimental Statistics. Courier Corporation (2013)
- 14. Gregoire, N.: and M.J.C.G.W.p. Bouillot, *Hausdorff distance* between convex polygons (1998)
- Mezhenin, A., Zhigalova, A.: Similarity analysis using Hausdorff metrics. in MICSECS. (2018)
- Williamson, D.F., Parker, R.A., Kendrick, J.S.: The box plot: A simple visual method to interpret data. Ann. Intern. Med. 110(11), 916–921 (1989)
- 17. Potter, K.C.: *Methods for Presenting Statistical Information: The Box Plot.* in *VLUDS.* (2006)
- Nex, F., Remondino, F.: UAV for 3D mapping applications: A review. Appl. Geomatics. 6(1), 1–15 (2014)
- Remondino, F., Fraser, C.: Digital camera calibration methods. Considerations and comparisons, in ISPRS Commission V Symposium 'Image Engineering and Vision Metrology', Dresden, Germany, September 25–27, 2006, H.-G. Maas and D. Schneider, Editors. ISPRS. p. 272. (2006)
- 20. Tang, R.: Mathematical Methods for Camera Self-Calibration in Photogrammetry and Computer Vision. (2013)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.