

Automatic Processing of Many Images for 2D/3D Modelling



Luigi Barazzetti, Marco Gianinetto and Marco Scaioni

Abstract The era of big data requires increasing automation for the analysis of huge information in a short time and this need becomes critical when dealing with geoinformation. This chapter describes the automatic geocoding of digital images based on high-end Photogrammetric and Remote Sensing methods. In particular, the so-called Structure-from-Motion (SfM) technique is developed to handle image data sets in close-range applications, and here, it is generalized to deal with multi-scale applications. Some examples are proposed with panoramic images for the measurement of indoor narrow spaces, with smartphone cameras and UAV for the 3D reconstruction of complex monuments, as well as with airborne and satellite images for the survey at the territorial scale.

Keywords Aerial · Archive photos · Close range · Drones · Satellites · Geocoding · Image registration · Photogrammetry · Remote sensing · Structure from motion

1 Fundamentals of Image Geocoding

Digital imagery is nowadays an enormous source of qualitative informative content, which may also provide accurate 2D and 3D metric information when they are processed using Photogrammetric and Remote Sensing techniques (Luhmann et al. 2014). In the past decades, the research in this field has developed powerful and consolidated methods for the automatic alignment of multiple data sets and the integration with the Image Processing and Computer Vision communities has played a paramount role. Today, the boundaries between these disciplines is much more blurred than it was in the past (Hartmann et al. 2016).

Disregarding the specific processes to extract 2D or 3D geometries from different types of images, *geocoding* (or *georeferencing*) is a fundamental phase to associate a

L. Barazzetti · M. Gianinetto (✉) · M. Scaioni
Architecture, Built Environment and Construction Engineering—ABC Department, Politecnico di Milano, Milan, Italy
e-mail: marco.gianinetto@polimi.it

© The Author(s) 2020

355

B. Daniotti et al. (eds.), *Digital Transformation of the Design, Construction and Management Processes of the Built Environment*, Research for Development, https://doi.org/10.1007/978-3-030-33570-0_32

location in the geographical space to each geometric primitive. In general, geocoding may be obtained from one out of the following approaches:

1. *Direct georeferencing* of the sensor is adopted for data acquisition by integrated Global Navigation Satellite Systems (GNSS), Inertial Navigation Systems (INS) and other types of sensors (e.g., star trackers in spaceborne sensors);
2. *Indirect georeferencing* is based on Ground Control Points (GCPs), whose coordinates are known in either the images space and the object space;
3. *Mixed techniques*.

When a time series made of images collected on the same location from the ground (e.g., fixed camera stations) or above the ground (e.g., drones, aircraft or satellite) is analysed, it is not necessary that all the frames are georeferenced. As in Photogrammetric *bundle block adjustment*, GCPs are measured on a smaller subset of the whole block of images, while the others are *co-registered* among them and with respect to it.

If *indirect georeferencing* techniques are used, one image is manually *geocoded* using known GCPs (this image is usually called ‘master’ or ‘reference’ image) and all the other images of the series are then manually or automatically co-registered to it. On the other hand, when using *direct georeferencing* techniques, all the images are already georeferenced and only a few GCPs are needed to correct some residual biases. Unfortunately, this approach is not possible for any type of applications, such as for instance, in close-range Photogrammetry (Luhmann et al. 2014). In other cases, it may only provide approximate geocoding to be used for instantiating other georeferencing techniques. This is the case of most Photogrammetric blocks recorded using drones (Colomina and Molina 2014; Granshaw 2018a) or for the analysis of satellite imagery, where direct geocoding is not enough accurate.

When indirect georeferencing methods should be called for, provided that some external constraints are always needed, the option to measure GCPs only on a (small) subset of the images and then co-register the rest of the data is really strategic to reduce the processing time and to limit the operator workload. Consequently, in recent years, several automatic approaches have been developed to this purpose. The Department of Architecture, Built Environment and Construction engineering (DABC) of Politecnico Milano has contributed to this topic by channelizing the registration process of different types of images within a common framework. This can be referred to the Photogrammetric procedure addressed to as *Structure from Motion*, which is discussed in the next section.

2 Structure from Motion

Structure from Motion (SfM) is, nowadays, the most popular technique used in Photogrammetry for image co-registration (in such a case, the correct term is *image orientation*). The reader can find a review of its origins and development in Granshaw (2018b), including the most relevant literature. In the beginning, SfM was limited to

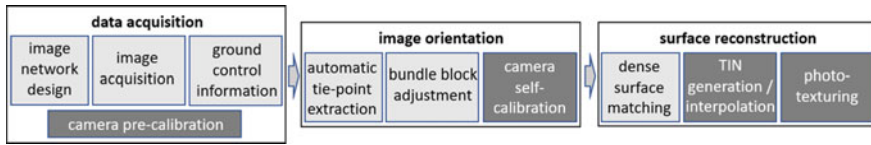


Fig. 1 Workflow of Structure-from-Motion basic processing pipeline

small–medium-size images (up to few megabytes) due to computational limitations of the implemented feature extraction algorithms such as SIFT or SURF (Barazzetti et al. 2009). Consequently, its use was limited to close-range blocks collected from ground-based stations or from drones.

Despite this, SfM was rapidly spreading out in the scientific community, where it reached a great popularity, especially in the field of Geosciences (Westoby et al. 2012; Eltner et al. 2016) and Cultural Heritage documentation (Barazzetti et al. 2011). The concurrent progress of the algorithms’ performances and the use of efficient computing solutions (e.g., parallel computing and processing at GPU level) has recently opened the way to deal with large blocks of large-size digital photos, such as the ones adopted in aerial projects for topographic mapping or Remote Sensing projects for land mapping. Figure 1 summarizes the basic concept of SfM, though many variations may come up in specific software implementations. Besides, the term SfM should only refer to the orientation stage, however, in the common language, it frequently refers to the whole Photogrammetric pipeline, also including the dense surface matching stage to generate a coloured point cloud.

2.1 Panoramic Images for the Survey of Indoor Narrow Spaces

Consumer-grade cameras able to capture 360° photos and videos are becoming more popular for the opportunity to look in any direction as well as for exploiting immersive visualization with virtual reality headsets (Barazzetti et al. 2018). Today, many cameras are already available on the market and some of them are quite cheap (100–500 EUR), such as the Ricoh Theta S, 360fly 4K, LG 360 CAM, Kodak PIXPRO SP360 4K, Insta360, Kodak PIXPRO SP360, or the Samsung Gear 360. On the other hand, professional systems for 360° imaging such as the GoPro Odyssey, SpheroCam V2, Nokia OZO or the GOPRO OMNI are still expensive today (dozens of thousands of euros).

As mentioned, 360° cameras capture the whole scene around the standpoint in a single shot and are becoming a new paradigm for Photogrammetry. Recalling that a 3D model of the scene can be created from images acquired from different points of view (at least two), multiple images can be processed following the typical workflow for image processing based on the spherical (equirectangular) camera model, whose equations in a camera centered reference system can be written as follows:

$$\begin{aligned} x &= f \arctan\left(\frac{X}{Y}\right) \\ y &= f \arctan\left(\frac{Y}{\sqrt{X^2+Z^2}}\right) \end{aligned} \tag{1}$$

where $f = \text{image_width}/(2\pi)$.

Two examples of Photogrammetric surveys using panoramic cameras are shown in Figs. 2 and 3. The first example describes the modelling of a narrow corridor (Fig. 2). While this is a conventional application for Photogrammetry, nevertheless, traditional Photogrammetric or (static) laser scanning surveys would have required a lot of time for the data acquisition because of the tight space. In this case, the collection and processing of more than 200 panoramic images required only few minutes.

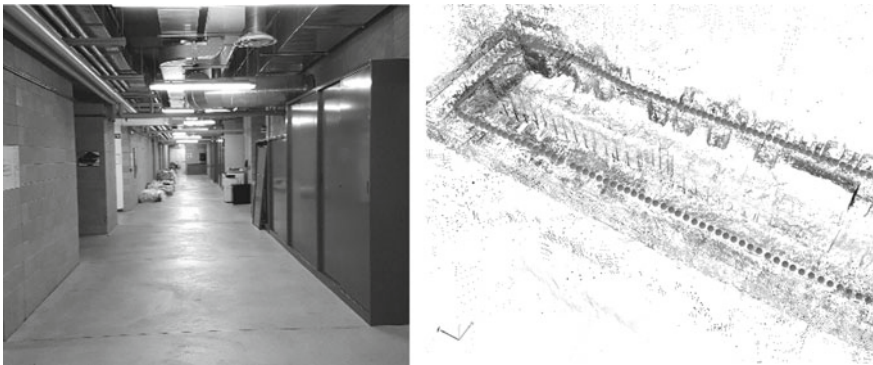


Fig. 2 Modelling of a narrow corridor with panoramic cameras

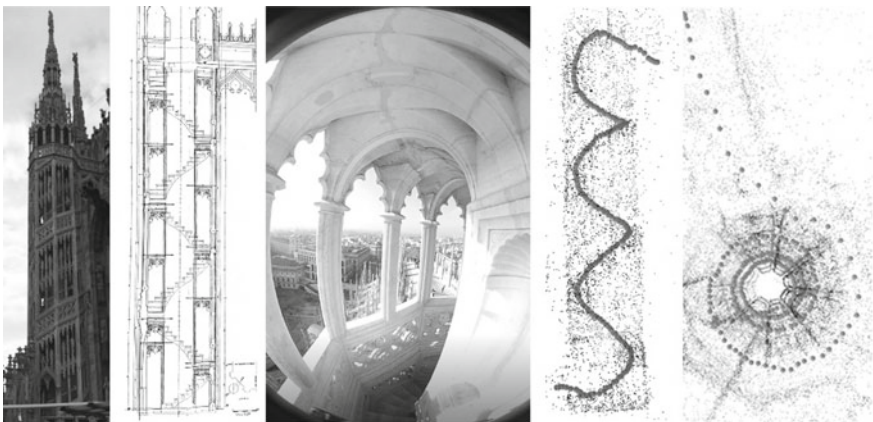


Fig. 3 Comparison of the shape of the staircase of the Pestagalli's spire in the Cathedral of Milan (on the left) with the model made from 360° panoramic images (on the right)

Figure 3 shows a non-conventional application. In this case, 360° panoramic images are used for surveying the narrow staircase inside the Pestagalli's spire of the Cathedral of Milan (Italy). Again, the tight space available makes spherical images a valid alternative to laser scanning or traditional Photogrammetric techniques.

These examples demonstrate that data processing is feasible not only for small images but also for large blocks. The geometry of a spherical image is more suitable than central perspective of traditional Photogrammetry when the field-of-view and image overlap are critical parameters. On the other hand, this solution may be problematic when considering very long image sequences, that could result in accuracy problems when images are progressively added without external constraints. In this case, the use of GCPs measured with a total station remains a primary tool to control the network geometry, as in the application examples abovementioned.

2.2 *Frame Images for the Survey of Complex Monuments*

If panoramic cameras are an emerging technology for 3D Photogrammetric surveys, frame cameras are a well-established technology for most applications. The SfM approach is an almost fully automated technique, nevertheless, in close-range projects, the image acquisition still remains under complete control of the operator. Yordanov et al. (2019) proposed some guidelines to drive this task, which is crucial to achieve final good results. Two examples with different types of photogrammetric blocks are presented to show how the same processing pipeline and the same software package (here Agisoft Metashape® ver. 1.5.0) may cope with the orientation of images based on SfM.

The first example (Fig. 4) describes the documentation of Cultural Heritage with a smartphone. In this case, 26 photos of a Medieval capital were collected during a visit to the 'Museum de los caminos' in the Episcopal Palace of Astorga (Spain). The internal camera of a Samsung Galaxy Grand Prime was used (focal length 3.3 mm, sensor size 3,264 × 1,836 pixels, pixel size 1.2 μm). Figure 4 (on the left)



Fig. 4 On the left: camera poses used to reconstruct a Medieval capital. On the right: the final textured 3D model of the capital obtained from the projection of images

shows the camera poses after image orientation based on SfM. Here, we can see that most photos were captured from stations in front of the capital and were located at approximately the same distance. A few images were taken with the camera rolled approximately 90° with respect to others in order to help camera calibration during the bundle adjustment applied to estimate image orientation parameters (Luhmann et al. 2016). An average number of 1,700 tie points were extracted per image, with a Root Mean Square (RMS) of residuals on reprojected image coordinates of 0.5 pixels.

Figure 4 (on the right) shows the textured 3D model obtained after dense matching (Remondino et al. 2014). Despite the lack of any quality assessment, the 3D model shows that very interesting results for documentation of a small Cultural Heritage element (the capital may be contained in a sphere of radius 60 cm) may be obtained using a cheap camera (about 100 EUR).

The second example concerns the reconstruction of an old industrial heritage building in Sicily (Italy). The ‘Fornace Penna’ was built up at Scicli (Ragusa province) at the beginning of XX Century, and then severely damaged by a fire in 1924, when it was abandoned. Figure 5 shows the camera poses after image orientation based on SfM and the textured 3D model obtained from the projection of images.

The complex building geometry and texture were surveyed using images from a small drone (DIJ Phantom 2, equipped with a camera featuring: focal length 3.61 mm, sensor size $4,000 \times 3,000$ pixels, pixel size $1.6 \mu\text{m}$) and some photos from ground-based stations. In total 252 images were recorded. This data set belongs to the SIFET benchmark described in Piras et al. (2017). Since the use of Unmanned Aerial Vehicles (UAVs) allows to predefine the acquisition points, thanks to the GNSS navigation system onboard, the geometry of camera poses can be carefully planned in advance (Pepe et al. 2018).

In such a case, the whole data acquisition was accomplished by a team of expert surveyors which also included the measurement of some GCPs for accurate georeferencing. These were realized by targets on the ground surface and natural features on the construction’s facades. Consequently, in such a case study also some evaluations

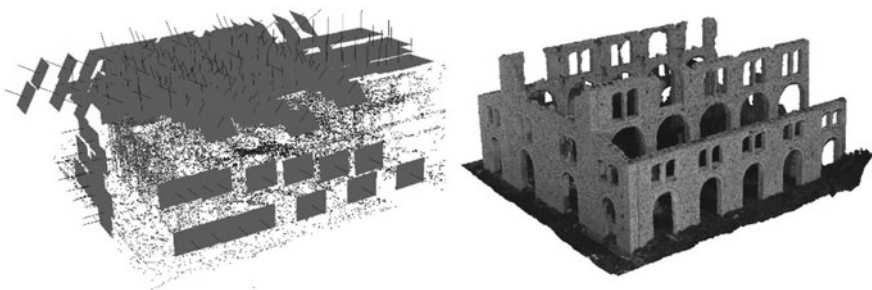


Fig. 5 On the left: camera poses of the UAV block of ‘Fornace Penna’ (Scicli, Italy). On the right: the final textured 3D model of the building obtained from the projection of images

of the accuracy were possible. An average number of 800 tie points were extracted per image, with an RMS of residuals on reprojected image coordinates of 0.4 pixels. Residuals on 23 GCPs resulted in an RMS error of 2.7 cm.

This second case study proves how the SfM approach combined with the use of UAV data allows modelling complex buildings in accurate and complete way.

2.3 Airborne and Satellite Images for the Survey at the Territorial Scale

When moving from the local to the territorial scale, aerial Photogrammetry and satellite Remote Sensing are the technologies used for extensive mapping.

Referring to aerial surveys, Fig. 6 shows a topographic mapping project over a small village in Italy. In this case, 18 images recorded with a ZI DMC-II digital airborne camera were processed using SfM (focal length 120 mm, sensor size 7,680 × 13,824 pixels, pixel size 12 μm). Digital airborne cameras are usually integrated into a GNSS/INS unit which may directly provide the exterior orientation parameters for georeferencing, besides the known interior orientation parameters that are used as input in the bundle adjustment. On the other hand, automatic aerial triangulation based on corresponding tie points observed on the images is applied to have more control over the final results. An average number of 1,400 tie points were extracted per image, with an RMS of residuals on reprojected image coordinates of 0.5 pixels. Here, the SfM approach was tested to replace the standard methods for automatic aerial triangulation adopted in aerial Photogrammetry.

In a similar way to the previous case studies, Fig. 6 shows some intermediate products of the SfM process. Though the example considered here only concerns a small block, this procedure may be also applied to larger blocks like the ones used in mapping projects at regional scales.

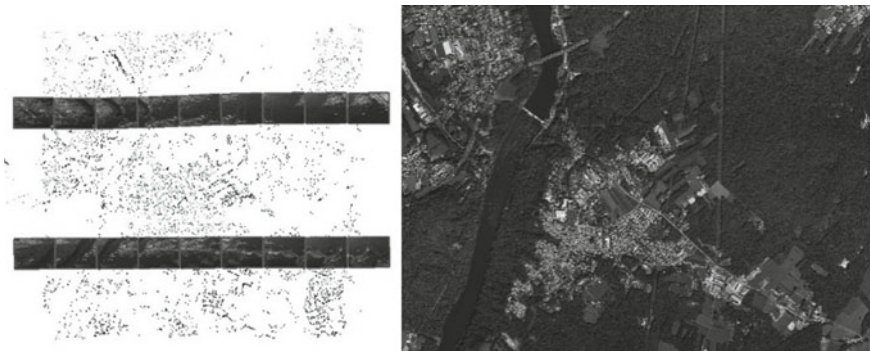


Fig. 6 On the left: camera poses used of the aerial block. On the right: the final orthophoto obtained from the projection of images on the obtained 3D model

When dealing with satellite Remote Sensing technologies, the exponential growth in the availability of traditional images and the revolution of microsatellites started with CubeSats pose several challenges. From the end-user point-of-view, commercial off-the-shelf software packages and cloud platforms able to process large number of images are becoming more accessible. As a consequence, new opportunities open up not only to experts but also to a wider public interested in digital survey and reconstruction at the territorial scale.

However, the improved availability of satellite data and products requires more efficient methods for data processing, in which the combined use of many images than the ones traditionally exploited is quite attractive. This task requires novel approaches and algorithms. Big data has, therefore, become a popular word in the Remote Sensing community and, at the same time, an opportunity and a challenge.

With reference to this matter, the Department ABC of Politecnico Milano has developed an alternative approach (MIRA—Multi-Image Robust Alignment) for the co-registration of large multitemporal data sets of images (Barazzetti et al. 2014) that extends the traditional pairwise approach (i.e., ‘one-to-one’) to the simultaneous processing of the whole time series (i.e., ‘one-to-many’). The basic concept is to use not only corresponding features between the ‘master’ image and the other images to be processed, but all the corresponding features shared between all the images of the time series. The coordinates of these correspondences (tie points) are used to instantiate a system of redundant equations to be solved within a Least Squares framework for the determination of the unknown registration parameters in a given geodetic datum. In such a way, also the images without corresponding features shared with the ‘master’ image can be registered, thus exploiting all the available images. In addition, this method increases the inner reliability of the observations and thus gain robustness against gross errors and limits the error propagation (Scaioni et al. 2018). Finally, with the growth of the time series, the network geometry improves, with overall benefits for the geometric alignment.

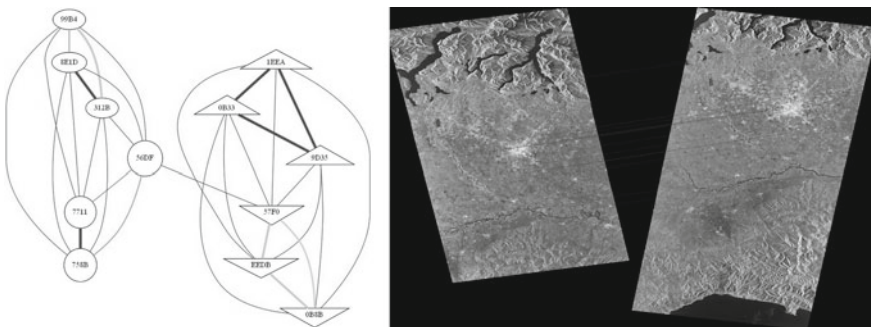


Fig. 7 On the left: connection graph of the Sentinel-1 SAR images. On the right: corresponding features between the ascending image (56DF) and the descending image (57F0) found by applying MIRA method

Figure 7 shows an example of data processing of Sentinel-1 (SAR) satellite time series with the MIRA method described above (Gianinetto et al. 2016). The surveyed region spans over an area of approximately 15,500 km² in the North-West of Italy, from the Lake of Como to the Gulf of Genoa. Twelve Sentinel-1 images (6 ascending and 6 descending) were collected in Stripmap mode (10 m geometric resolution), with dual polarization (HH/HV for ascending, VV/VH for descending), and with large differences in incidence viewing angles. This is a very challenging acquisition scheme for image registration.

Looking at the connection graph (Fig. 7 on the left) we can see that some images have a strong interconnection, while some other images have a weak or null interconnection. That reflects the imaging modes, in fact: (i) elliptical edge nodes refer to S2 ascending images; (ii) circular edge nodes refer to S4 ascending images; (iii) triangular edge nodes refer to S1 descending images; and (iv) inverse triangular edge nodes refer to S6 descending images. In this case, traditional pairwise methods for image co-registration failed to process this data set because many images did not have any correspondences with the ‘master’ image. On the other hand, the simultaneous processing based on MIRA exploited the link between the ascending and the descending blocks of images (56DF vs. 57F0) along with the links between each image pair to calculate a global co-registration of the entire satellite time series.

3 Conclusions

The growing availability of digital images captured with different sensors is an amazing opportunity for several users involved in many real applications.

When the images must be turned into metric products, the use of automated, efficient and reliable registration procedures has a primary importance to prepare all the data for the following stages of processing. Such preliminary operation cannot be neglected when accurate deliverables must be produced, requiring robust solutions able to deal with huge datasets in a fully automated way.

The integration of the SfM strategy in Photogrammetry and Remote Sensing has led to novel processing algorithms, which have significantly changed the traditional workflow of several surveying applications based on digital images.

Acknowledgements The authors would like to thank the Italian Society of Photogrammetry and Topography (SIFET) for the UAV benchmarking data set ‘Fornace Penna’.

References

Barazzetti L, Remondino F, Scaioni M (2009) Combined use of photogrammetric and computer vision techniques for fully automated and accurate 3D modeling of terrestrial objects. In: Video-metrics, range imaging, and applications X, SPIE, paper no. 74470M

- Barazzetti L, Binda ML, Scaioni M, Taranto P (2011) Photogrammetric survey of complex geometries with low-cost software: application to the 'G1' Temple in Myson, Vietnam. *J Cult Herit* 12:253–262
- Barazzetti L, Gianinetto M, Scaioni M (2014) Automatic co-registration of satellite time series via Least Squares adjustment. *Eur J Remote Sens* 47:55–74
- Barazzetti L, Previtali M, Roncoroni F (2018) Can we use low-cost 360 degree cameras to create accurate 3D models? *Int Arch Photogramm Remote Sens Spat Inf Sci* 42(2):69–75
- Colomina I, Molina P (2014) Unmanned aerial systems for photogrammetry and remote sensing: a review. *ISPRS J Photogramm Remote Sens* 92:79–97
- Eltner A, Kaiser A, Castillo C, Rock G, Neugirg F, Abellán A (2016) Image-based surface reconstruction in geomorphometry—merits, limits and developments. *Earth Surf Dyn* 4:359–389
- Gianinetto M, Monno V, Barazzetti L, Dini L, Daraio MG, Rota Nodari F (2016) Subpixel geocoding of COSMO-SkyMed and Sentinel-1 time series imaged with different geometry. In: International geoscience and remote sensing symposium (IGARSS 2016), Beijing, China, pp 5007–5010
- Granshaw SI (2018a) RPV, UAV, UAS, RPAS ... or just drone? *Photogramm Rec* 33:160–170
- Granshaw SI (2018b) Structure from motion: origins and originality. *Photogramm Rec* 33:6–10
- Hartmann W, Havlena M, Schindler K (2016) Recent developments in large-scale tie-point matching. *ISPRS J Photogramm Remote Sens* 115:47–62
- Luhmann T, Robson S, Kyle S, Boehm J (2014) Close-range photogrammetry and 3D imaging. Walter de Gruyter, p 684
- Luhmann T, Fraser CS, Maas H-G (2016) Sensor modelling and camera calibration for close range photogrammetry. *ISPRS J Photogramm Remote Sens* 115:37–46
- Pepe M, Fregonese L, Scaioni M (2018) Planning airborne photogrammetry and remote-sensing missions with modern platforms and sensors. *Eur J Remote Sens* 51:412–435
- Piras M, Di Pietra V, Visintini D (2017) 3D Modeling of industrial heritage building using COTSs system: test, limits and performances. *Int Arch Photogramm Remote Sens Spat Inf Sci* 42(2/W6), 281–288
- Remondino F, Spera MG, Nocerino E, Menna F, Nex F (2014) State of the art in high density image matching. *Photogramm Rec* 29:144–166
- Scaioni M, Barazzetti L, Gianinetto M (2018) Multi-image robust alignment of medium-resolution satellite imagery. *Remote Sens* 12(24)
- Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds J (2012) 'Structure-from-Motion' photogrammetry: a low-cost, effective tool for Geoscience applications. *Geomorphology* 179:300–314
- Yordanov V, Mostafavi A, Scaioni M (2019) Distance-training for image-based 3D modelling of archaeological sites in remote regions. *Int Arch Photogramm Remote Sens Spat Inf Sci* 42(2/W11):1165–1172

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.

