Architectural Heritage: 3D Documentation and Structural Monitoring Using UAV*

Danila Germanese[1✉], Maria Antonietta Pascali[1], Andrea Berton[2], Giuseppe Riccardo Leone[1], Davide Moroni[1], Bushra Jalil[1], Marco Tampucci[1], and Antonio Benassi[1]

[1] Institute of Information Science and Technologies - CNR, via G. Moruzzi 1, Pisa 56127, Italy,
danila.germanese@isti.cnr.it
[2] Institute of Clinical Physiology - CNR, via G. Moruzzi 1, Pisa 56127, Italy

Abstract. Architectural heritage preservation and dissemination is a very important topic in Cultural Heritage. Since ancient structures may present areas which are dangerous or difficult to access, Unmanned Aerial Vehicles may be a smart solution for the safe and fast data acquisition. In this paper we propose a method for the long term monitoring of cracking patterns, based on image processing and marker-based technique. Also the paper includes the description of a pipeline for the reconstruction of interactive 3D scene of the historic structure to disseminate the acquired data, to provide the general public with info regarding the structural health of the structure, and possibly to support the drone pilot during the survey. The Introduction provides a state of the art about the crack monitoring from visible images; it follows a description of the proposed method, and the results of the experimentation carried out in a real case study (the Ancient Fortress in Livorno, Italy). A specific section is devoted to the description of the front-end of augmented reality designed for heritage dissemination and to support the drone usage. Details about the future works conclude the paper.

Keywords: Crack quantification · Crack monitoring · Photogrammetry · UAV · 3D rendering.

1 Introduction

The structural deterioration of architectural heritage is an old problem. Among the large set of structural features of an ancient building or structure, we devoted this paper to the problem of monitoring and measuring missing or deformed structural elements, cracks and fissures. Nowadays, visual inspection is the most used technique to detect damage or to evaluate their variation over time. Nonetheless, such technique may be time consuming and expensive, and even not possible if the access to critical locations is forbidden for safety reasons.

* Partially supported by the Tuscany Region, FAS program.

On the other hand, the constant growth of digital technologies, led to novel and efficient low-cost hardware and applications, supporting the monitoring of specific regions and the assessment of the mechanical stability, thus preventing critical events.

In this paper we describe a method to track over time the variations of the cracking patterns in buildings. Also we show its applicability in a challenging case study, namely the *Fortezza Vecchia* – an ancient fortress in Livorno (Italy). The study was carried out in the framework of MOSCARDO project [9], devoted to the development of a monitoring system able to collect and process data about the structural health of ancient buildings, and to provide alert notifications in case of anomalies; of course, the implementation of such a functionality requires the acquisition of a reference data set over at least one year, in order to set reliable threshold values, e.g. with respect to the seasonal variations (vegetation, environmental light, temperature, humidity). Among the techniques aiming at monitoring the structural integrity of buildings, we used the marker-based method, as it allows for a non-destructive analysis of cracking patterns. We also demonstrated, in a previous pre-application study [8], the feasibility of this approach.

2 State of the art

The literature provides a number of studies proposing alternative methods to visual inspection. Such methods, in general, divide into two different groups: invasive and non-invasive approaches. Here, we briefly list a comment some of the non-invasive methods, in particular those which are more suitable to be applied in cultural heritage.

Close-range digital photogrammetry includes a large family of methods, as reported by Remondino and co-workers [17]. In general, it is based on several acquisitions of a set of images used to produce a 3D point cloud of the scene. To monitor over time the crack opening, the point clouds generated at different dates are compared. Such comparison may be performed by exploiting different techniques, such as (i) conventional analysis, which uses statistical tests to compare the estimated 3D coordinates of the same points [22]; (ii) shape analysis, e.g. by matching surfaces [10] or comparing their shape signatures [3], or comparing a specific shape parameter (the surface area associated to each crack) complemented with a bootstrap testing to detect only statistical meaningful variations in cracking pattern [1]. However, among these papers, only the last one tackles the application of 3D shape methods to the crack analysis, representing the crack as a surface, hence not allowing for the monitoring of specific critical points along the crack.

Other methods aim at automatically detecting and measuring structure damages and cracks using image-based algorithms, which allow for specifically filtering out the cracking patterns, such as in [5]. In this work, two pipelines for the crack segmentation are described: the former evaluates the color level for each pixel, and enhance the structural discontinuities by adding more "white" or more ”black” to make the structural discontinuities even darker (unfortunately this approach fails when the structure walls are not clear); in the second one, the cracking pattern is filtered out by automatically detecting edges (by applying a Gaussian blur and subtracting the filtered image from the original one). Even if this study provides image-based techniques for the crack segmentation no quantitative analysis of the structural damages is proposed.

The method proposed by Jahanshahi et al [13], [12] is based on 3D reconstruction of the scenario, image segmentation and binarization, and two classifiers (SVN and NN) to detect, isolate and distinguish the pattern related to small cross-sectional structural defects (0.4-1.4mm). On one hand, such method showed robustness when applied to images captured from any distance (20 m in their experimental tests) and acquired using any resolution and focal length (600 mm in their experimental tests); on the other hand, it seemed to be suitable only for detection of small cross-sectional defects over homogeneous background.

In general, the principal drawbacks that image-based methods show are related to noise removal, edge detection, registration, application of morphological functions, colour analysis, texture detection, segmentation. Among the major challenges, the removal of noise due to the edges of doors, windows, and buildings. In the work of Jahanshahi [11] the pro and cons of cracks automatic detection in civil infrastructures performed with image-based methods is discussed.

Ellenberg [4] describes the main issues related to image acquisition performed by UAV, i.e. the environmental conditions, the setting of camera parameters, the distance of the camera (the greater the distance of the camera, the lower the accuracy with which the crack width will be calculated), and angle of orientation.

Other studies aimed at marking the most critical points of a crack: any displacement identified by the coordinates of the targets is used to calculate the force field (tensile and shear forces) along the discontinuities of interest. Such approach aim at enhancing the accuracy and the repeatability of the feature point extraction task [21].

Nishiyama and colleagues [16] exploited reflective targets (i.e., targets made by glass droplets, in order to reflect the light as much as possible) to mark the point of interest of a discontinuity. They aimed at assessing the displacements of the two surface portions of the crack by acquiring a number of images of the marked crack, implementing photogrammetric algorithms and calculating the coordinates of the targets.

In [21], another method, based on Hough transform and homography techniques, is proposed to correct the perspective error, detect targets, and identify their planar coordinates and geometric centers.

Benning et al. [2] prepared the surface of structural elements of pre-stressed, reinforced and textile concrete by a grid of circular targets. Images were captured simultaneously and the measure of the relative distances between adjacent targets was repeated in time intervals. Therefore, the evolution of the cracks and discontinuities present on the surface of the concrete specimens was monitored over time. In addition, a Finite-Element-Module was developed, which simulated the test: thus, the results of photogrammetric measurements were compared with the numeric tension calculation and iteratively improved. In these studies, the method was applied only in the case of planar, small cross-sectional cracks.

The markers can also be home-made, as reported in [19], in combination with useful suggestions on their dimensions, materials. Such method suffer from the position of the camera: the greater the distance of the camera, the lower the accuracy with which the centroid coordinates will be calculated.

Properly designed for cultural heritage, our solution addresses the major open problems related to the cracking pattern of the ancient structures, where the fissures may be wide, large, and often non-planar. The markers allow for an accurate quantitative measure and for a long-term monitoring. In addition, the use of the UAV allows for exploring all the most critical points of the structure, even the non-accessible ones.

3 Methodology

Our main case study is the Fortezza Vecchia, an ancient fortress in Livorno (Italy): its walls are difficult to be monitored, because the fortress is partially surrounded by the sea (see Fig. 1). A very important feature of most of the cracks along the walls of the fortress is that their sides are quite far and definitely they don't lie on the same plane, as visible in Fig. 1, Bastione della Capitana. It makes very difficult to obtain an absolute and accurate measurement of the separation of the sides with standard methods. For this kind of structures, irregular, outdoor, subject to environmental agents and to seasonal changes (for vegetation or even some weeds on it), photogrammetric reconstruction quality may be not enough, especially when the aim is to compare over time measurements of the cracks along the structure. Hence, we decided to use markers to provide a complete and stable 3D information about specific fiducial points along the crack, to be tracked over time. Such data include: *(i)* the set of the 3D coordinates of each marker's corners, *(ii)* the set of the distances between the barycenters of each pair of markers, and *(iii)* the angle variations between the reference frame associated to each marker. Other advantages of such minimally invasive technique are that it enables:

- reaching high accuracy when performing a quantitative analysis of the crack;
- using Unmanned Aerial Vehicles (UAVs) to acquire and, possibly, process on board the images:

Also, the markers represent optimal reference points with respect to the use of UAV for data acquisition. It is worth noting that UAV-based technologies allow for a fast and highly repeatable data acquisition, even in area difficult to access, thus reducing costs and risks.

We chose to use the ArUco markers because they are black and white square planar coded markers [6] easy to use, and they can be reliably detected under a wide range of environmental conditions; in addition, such markers allow for an

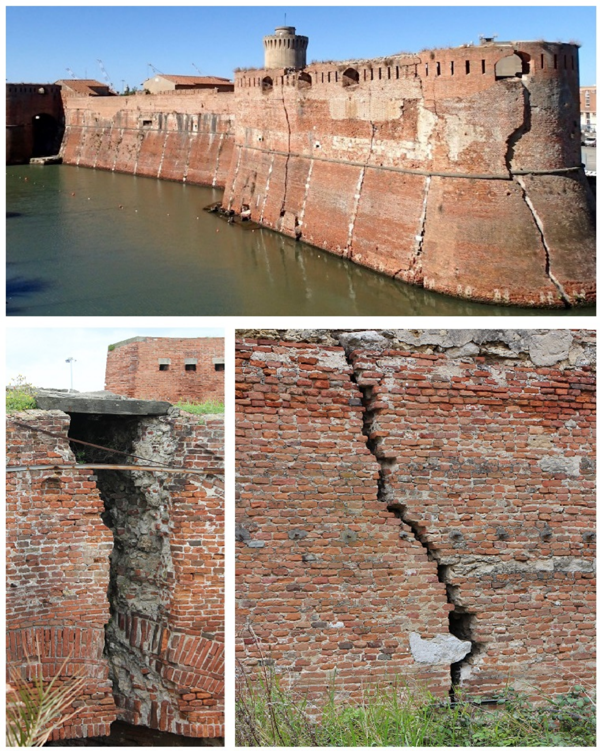

Fig. 1. Evident structural defects in the walls of the ancient fortress, in the area named *Bastione della Capitana* (Fortezza Vecchia, Livorno, Tuscany, Italy).

accurate and robust camera localization. Our method is based on the Simultaneous Localization and Mapping, described in [15], which is optimized for the creation of 3D map of the markers visible in the images acquired. A sequence of frames of the same scene is acquired and at each frame the graph-pose is estimated minimizing the re-projection error in the detection of the marker corners. The output of the algorithm are the 3D coordinates of the corners with marker ids. Then, the Euclidean distances between the markers' barycenters and poses are computed.

In the previous pre-application work [8], we showed the feasibility of such a marker-based approach for the assessment of the crack opening, simulated in a laboratory. We estimated a measurement accuracy of less than 1 mm, using a 18 Mega-pixels mirror-less (Canon EOS M), with a focal length set at 24 mm. The acquired images have a resolution of 5184 x 3456 pixels. The same hardware has been used for the experimentation described in the Section 4. As regards the visual impact of the markers placed on an ancient structure, such as the wall of the Old Fortress in Livorno, we printed markers with two different side lengths: 0.1 m and 0.2 m. We found that, as all of them were detected correctly , the smaller ones may be used for the next experimentation, without loosing in accuracy.

4 Experimental setting

Six pairs of markers, four small pairs (0.1 m side length) and two big ones (0.2 m side length), have been glued along to the sides of a complete vertical cut in the walls of Fortezza Vecchia in Livorno, located in the area called *Bastione della Capitana*: the highest pair of markers is glued to the wall about 3.3 m from the ground. The camera used for 3D reconstruction and photogrammetry is the Canon EOS M, a 18 Mega-pixel mirror-less with a sensor APS-C of 22.3 x 15 mm (aspect ratio 3 : 2). The maximum video resolution is of 1920 x 1080 pixel at 30 fps. It weighs 298 g and has dimensions 108 x 66 x 32 mm. The focal length varies in the range 18-55 mm. E.g., setting the focal length at 24mm, and the target at 1.5 mt, the field of view will be of 1.39 m (width) and 0.93 m (height), and the pixel resolution (computed from the camera fact-sheet) will be of 0.27 mm. The camera has been calibrated using a ChArUco board as in [8], and at each acquisition a sequence of at least 6 images were acquired 3 m, 6.5 m and 9 m far from the wall. The focal length of the camera has been set at 24 mm; this implies that, for example, when the camera is 6 m far from the target, the field of view is of 5.575 m (width) and 3.725 m (height), and the pixel resolution (computed from the camera fact-sheet) is of 1 mm.

Another set of data have been collected recently, but not yet processed, surveying the area with a drone. The drone is a Micro Air Vehicle designed and assembled at the Institute of Information Science and Technologies of the National Research Council of Italy. Having a drone following a predefined path (made of GPS waypoints), allows to repeat the same acquisition; hence supporting the creation of a large dataset of the site of interest over time.

5 Results

As shown in Fig. 2 all markers are correctly detected, and the corners' coordinates are used to compute the distances between the barycenters, to roughly check the correlation with the measures performed with a flexible meter (see Table 1). Being this lesion large and not planar, linear measurements on it are inherently not accurate.

Table 1. Barycenter distances and approximate ground truth. Last two columns show the absolute value of the difference between the approximate ground truth and the barycenter distances computed from the two acquisitions. All values are in m.

Marker pair (ids)	~Ground Truth	1^{st} acq.	2^{nd} acq.
id: 27-20	0.33	0.001	0.000
id: 38-23	0.385	0.004	0.003
id: 3-4	0.55	0.007	0.007
id: 31-26	0.33	0.009	0.012
id: 25-18	0.42	0.003	0.011
id: 5-10	0.51	n.a.	0.006

Table 2. Barycenter distances, first and second acquisition, all values are in m.

Marker pair	1^{st} acq.			2^{nd} acq.		
ids	3 m	6.5 m	9 m	3 m	6.5 m	9 m
27-20	0.329	0.333	n.a.	0.332	0.333	0.336
38-23	0.386	0.391	n.a.	0.387	0.390	0.394
3-4	0.555	0.557	0.562	0.552	0.558	0.559
31-26	0.334	0.338	0.346	0.334	0.336	0.341
25-18	0.428	0.434	0.444	0.428	0.430	0.436
5-10	0.522	0.523	0.526	0.540	0.526	0.522

In Table 2, the distances between the markers' barycenters are computed for the two sets of images manually acquired 3 m, 6.5 m and 9 m far from the crack. As expected, the approximately ground truth is not always close enough to the computed distances; nonetheless, the two sets of values from the two acquisitions seem to agree enough to justify further efforts and more experimentation to refine the method (also regarding the acquisition procedure and the hardware setting) in order to increase its accuracy in the outdoor setting; for so large cracks in brick ancient structure, we aim at an accuracy of 1 mm, considered a meaningful measure. Table 3 shows that in our tests the accuracy of the measurement, even if showing a bit large standard deviation, does not depend on the marker size. On the other hand, Table 2 shows that the stability of the measurement should

Fig. 2. ArUco fiducial markers have been placed along the crack in the most critical points. The markers are correctly detected and identified.

be increased with respect to the camera-target distance. This last point is quite important with respect to the usage of UAV for the data acquisition.

Table 3. Chart of average values and standard deviation computed with respect to the two acquisitions, not regarding the distance camera-target (all values are in m).

Marker pair	average	st dev
id: 27-20	0.332	0.002
id: 38-23	0.39	0.003
id: 3-4	0.557	0.004
id: 31-26	0.338	0.004
id: 25-18	0.433	0.006
id: 5-10	0.527	0.007

This results will be enriched by the analysis of the 3D pose of each marker, made through the computation of the angle variations between the reference frame associated to each marker.

6 Augmented reality

In the management of the cultural heritage, 3D rendering, recording and documentation are a fundamental step [7] towards the enhancement of cultural heritage. General public may experience an interactive survey of the ancient structures, possibly including not accessible areas, and explore all the digital information extracted. As reported in [18], several technologies can be used to build a 3D digital model. LiDAR (Light Detection And Ranging) is a high quality remote sensing method that uses light in the form of a pulsed laser to measure ranges. Despite the proven quality of the LiDAR systems, we preferred to explore less expensive solutions that achieve comparable results. A simple and lightweight monocular camera can be used for Structure from motion (SfM), in order to build the 3D model of the structure from a sequence of different views of the object. The number, the quality, and the resolution of the images have to be taken into account because they may affect very much the time needed to obtain a full 3D reconstruction.

SfM algorithm, in most of its implementations, consists of the following main steps: *(i)* feature point selection in all the images, generally obtained by applying SIFT (scale-invariant feature transform) or SURF (speeded-up robust features), *(ii)* matching of the corresponding features and registration between images (incorrect matches are usually filtered out with specific algorithms, e.g. RANSAC, random sample consensus), *(iii)* detection of the control points, *(iv)* building of a dense point-cloud, that is performed by using wide baseline stereo correspondence [20], and finally *(v)* the surface reconstruction as a polygonal mesh, that is the final 3D model of the object.

In general, 3D reconstruction algorithms may take hours or days using a normal pc. The using of multiple parallel GPUs or of cloud computing are strongly recommended to speed up the whole process. The most popular software of 3D reconstruction include **Agisoft Photoscan**[3], which is the first photogrammetric software and allows the user tuning the reconstruction parameters during the procedure to increase the quality, depending on the input data, on the user preferences, and on the computing resources; **COLMAP**[4], an open-source software which allows for setting the reconstruction parameters, but does not allow for interacting with the middle result of the reconstruction phases; **Autodesk Recap Photo**[5], which exploits cloud technologies, it is able to processes up to 100 photos at once, but it does not allow for setting any parameter relative to the reconstruction phases.

Beyond computing dense detailed models [14], we got the best result from Agisoft Photoscan, by selecting the appropriate key points and processing only the interesting part of the image. Then, a virtual scene containing the reconstructed object has been created by using the Unity[6] engine. The exploitation of such type of engine guarantees the easiness of navigation and, at the same time, the overall representation quality. Inside the scene, users can easily navigate around the reconstructed object and have a quick-look of all the regions of interest of the structure, e.g. sensors installed to monitor environmental or structural parameters, as shown in Fig. 3. Cracks may be highlighted and labelled with latest measurements. It is also possible to interact with the cracks

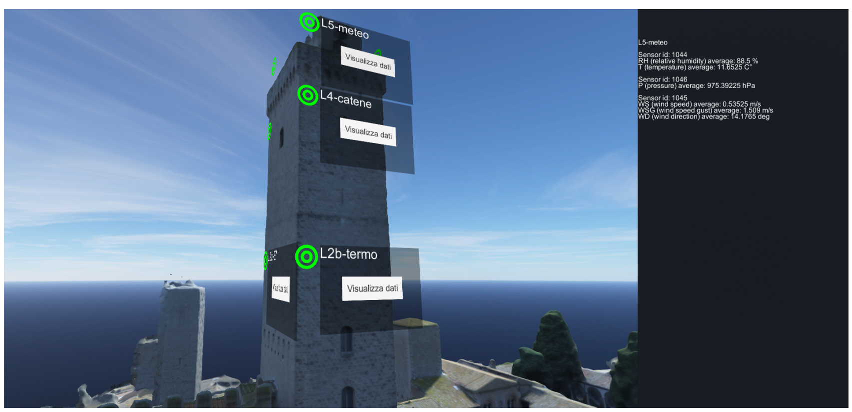

Fig. 3. 3D front end interface. In the main panel, the 3D reconstruction of Torre Grossa in San Gimignano, Italy, can be explored. Installed sensors are gathered from MOSCARDO database and displayed directly. When a sensor is selected, last retrieved data are shown on the right panel.

[3] www.agisoft.com
[4] colmap.github.io
[5] www.autodesk.com/products/recap
[6] www.unity3d.com

in order to retrieve past calculated values or visualize charts representing the crack opening evolution over time. The 3D reconstruction of the scene is also useful to support the drone survey when an alert notifications in case of critical anomalies is provided.

7 Conclusions and future works

Currently only a few acquisitions have been performed and the plausibility of the proposed method has been confirmed. Since the preliminary results are promising, further effort will be devoted to: acquire and process more data acquired temporally close to each other in order to assess the robustness of the measurement method; acquire and process data on a monthly basis, for at least one year, to track the evolution of the crack features and possibly develop a model to detect critical trends; and to test the validity of the UAV application.

Also, our system is minimally invasive, as required in the field of cultural heritage, the measurements provided are accurate is not enough, it opens the way to further structural analysis based on the 3D geometric representation of a structural defect, e.g. looking not only at all the barycenter distances but also at the pose variation of each marker. Note that our approach works around the main difficulties of a crack segmentation, in our specific scenario: outdoor, large and non planar cracks of ancient structures.

On the other hand, the 3D information and technology reveals to be quite a powerful tool: the reconstruction of the architectural asset could be used to disseminate the cultural heritage to the general public, and even may support the expert to analyse the structural data acquired by sensors or the pilot of the drone to flight over the specific area interested by a detected anomaly.

8 Acknowledgements

This work has been carried out in the framework of the project MOSCARDO (FAR-FAS 2014) and SCIADRO (FAR-FAS 2014).

References

1. Armesto, J., Arias, P., Roca, J., Lorenzo, H.: Monitoring and assessing structural damage in historical buildings. The Photogrammetric Record **21**, 269–291 (2006)
2. Benning, W., Görtz, S., Lange, J., Schwermann, R., Chudoba, R.: Development of an algorithm for automatic analysis of deformation of reinforced concrete structures using photogrammetry. VDI Berichte **1757**, 411–418 (2003)
3. Cardone, A., Gupta, S., Karnik, M.: A survey of shape similarity assessment algorithms for product desing and manifacturing applications. Journal of Computing and Information Science in Engineering **3**, 109–118 (2003)
4. Ellenberg, A., Kontsos, A., Bartoli, I., Pradhan, A.: Masonry crack detection application of an unmanned aerial vehicle. In: Proceedings of Computing in Civil and Building Engineering, ASCE2014 (2014)

5. Eschmann, C., Kuo, C., Kuo, C., Boller, C.: Unmanned aircraft systems for remote building inspection and monitoring. In: 6th European workshop on structural health monitoring (2012)
6. Garrido-Jurado, S., Munoz-Salinas, R., Madrid-Cuevas, F., Marín-Jiménez, M.: Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition **47**(6), 2280 – 2292 (2014). https://doi.org/http://dx.doi.org/10.1016/j.patcog.2014.01.005
7. Georgopoulos, A., Stathopoulou, E.K.: Data acquisition for 3D geometric recording: state of the art and recent innovations, pp. 1–26. Springer (2017)
8. Germanese, D., Leone, G.R., Moroni, D., Pascali, M.A., Tampucci, M.: Long-term monitoring of crack patterns in historic structures using uavs and planar markers: A preliminary study. J. Imaging **4**(99) (2018)
9. Germanese, D., Leone, G.R., Moroni, D., Pascali, M.A., Tampucci, M.: Towards structural monitoring and 3d documentation of architectural heritage using UAV. In: Multimedia and Network Information Systems - Proceedings of the 11th International Conference MISSI 2018, Wrocław, Poland, 12-14 September 2018. pp. 332–342 (2018). https://doi.org/10.1007/978-3-319-98678-4_34
10. Gruen, A., Akca, D.: Least square 3d surface and curve matching. ISPRS Journal of Photogrammetry & Remote Sensing **59**, 151–174 (2005)
11. Jahanshahi, M.R., Kelly, J.S., Masri, S.F., Sukhatme, G.S.: A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures. Structure and Infrastructure Engineering **5**(6), 455–486 (2009)
12. Jahanshahi, M.R., Masri, S.F.: A new methodology for non-contact accurate crack width measurement through photogrammetry for automated structural safety evaluation. Smart materials and structures **22**(035019) (2013)
13. Jahanshahi, M.R., Masri, S.F., Padgett, C.W., Sukhatme, G.S.: An innovative methodology for detection and quantification of cracks through incorporation of depth perception. Machine vision and applications **24**(2), 227–241 (2011)
14. Majdik, A.L., Tizedes, L., Bartus, M., Sziranyi, T.: Photogrammetric 3d reconstruction of the old slaughterhouse in budapest. In: Computational Intelligence for Multimedia Understanding (IWCIM), 2016 International Workshop on (2016)
15. Munoz-Salinas, R., Marin-Jimenez, M., Yeguas-Bolivar, E., Medina-Carnicer, R.: Mapping and localization from planar markers. P. Recog. **73**, 158–171 (2018)
16. Nishiyama, S., Minakata, N., Kikuchi, T., Yano, T.: Improved digital photogrammetry technique for crack monitoring. Adv. Eng. Informatics **29**(4), 851–858 (2015)
17. Remondino, F., El-Hakim, S.: Image-based 3d modelling: a review. The Photogrammetric Record **21**, 269–291 (2006)
18. Remondino, F., Campana, S.: 3D Recording and Modelling in Archaeology and Cultural Heritage - Theory and Best Practices. Archaeopress BAR (2014)
19. Shortis, M.R., Seager, J.W.: A practical target recognition system for close range photogrammetry. The Photogrammetric Record **29**(147), 337–355 (2014)
20. Strecha, C., Fransens, R., Gool, L.V.: Wide-baseline stereo from multiple views: A probabilistic account. In: In CVPR. pp. 552–559 (2004)
21. Valença, J., Dias-da Costa, D., Júlio, E., Araújo, H., Costa, H.: Automatic crack monitoring using photogrammetry and image processing. Measurement **46**(1), 433–441 (2013)
22. Welsch, W., Heunecke, O.: Models and terminology for the analysis of geodetic monitoring observations. In: FIG 10th International Symposium on Deformation Measurements. International Federation of Surveyors. vol. 25, p. 22 (2001)