The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-M-2-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

# PRACTICAL IMPLEMENTATION OF PHOTOGRAMMETRY FOR THE MODELLING OF A CYLINDRICAL HISTORICAL BUILDING

E. Simonetto [1] *, C. Charlet [1], E. Labergerie [1], G. Batifol[1], T. Guivarch[1], T. Le Goff[1], C. Senra[1]

[1] GeF, Cnam-ESGT, Le Mans, France- (elisabeth.simonetto, christophe.charlet2)@lecnam.net

**KEY WORDS:** Close-range photogrammetry, image capture, strategy, cylindrical structure, 3D point cloud, orthomosaic, cultural heritage.

**ABSTRACT :**
This work deals with the image capture for close-range photogrammetry in the context of the modelling of a round tower with repetitive texture (made of bricks), both inside and outside. For that purpose we test different acquisition strategies that differ in terms of camera path (straight or circular) and number of images acquired at each camera position, that is with or without oblique images. Besides, the formula to compute the curvilinear base according to a given overlap rate between successive frontal photos are written for each strategy. The comparisons rely on the observation of the aspect of the different dense point clouds (noise, holes), the deformation seen in the orthomosaics and values extracted from the photogrammetric projects such as metric accuracies with GCP. Our results confirm some art rules.

## 1. OBJECTIVES

Nowadays, photogrammetry and laser scanning technique are common way to model the cultural heritage. The choice of one or both techniques depends on the object nature and the objectives of the measurements (Boehler & Marbs, 2004, Pavlidis et al., 2007, Grussemeyer et al., 2008).
Using photogrammetry, the SfM (Structure from Motion) approach and MVS (Multi-View Stereo) algorithm lead to a dense 3D point cloud. However, one has to acquire the images in an appropriate way to avoid holes, irregular point density and noise effects. This may all the more occur in case of complex shapes (involving occlusions), particular materials (for instance, with a repetitive or homogeneous texture), and some environmental conditions (lighting, obstacles).
In this work, we test some practical implementations for the photogrammetric capture of a round building such as an historical brick or stone tower. Interior faces 3D modelling is considered as well as the exterior one.

## 2. CONTEXT

Recommendations are already well known to acquire suited images for a photogrammetric project. One must ensure enough redundancy to produce a robust assessment of the camera parameters and a dense 3D point cloud. They come from art rules and traditional photogrammetry, such as the CIPA 3x3 rules derived from (Waldhäusl & Ogleby, 1994) or the work presented in (Wenzel et al., 2013) and inspired from (Kraus, 2007). Other works have proposed image datasets. Some are acquired owing to a specific instrumentation such as a camera arm and others have been simulated. These datasets allow to test different photogrammetric methods as done in (Ahmadabadian et al., 2013), (Marelli et al., 2020) or (Caudal, 2021).

Thus, to model a planar facade by terrestrial photogrammetry, one moves along the wall to a given distance, taking photos with a panoramic strategy at each station (that is one camera position). The baseline, distance between each station, is determined in order to have a good overlap between successive photos (around 80%). This overlap should allow to see every points in at least three images (Tuttas et al., 2016). The capture distance from the wall is chosen so that the researched ground pixel size could be achieved, which is linked to the desired level of details:

$$GSD = t\frac{D}{f} \quad (1)$$

where $GSD$ = ground sampling distance or ground pixel size
$D$ = shooting distance
$t$ = image pixel size
$f$ = focal length

Nevertheless, the base to distance ratio should be chosen so that the angle between the camera axis from two successive views is less than 15° on the object (Garcia Gago et al., 2014).

When the faces are round such as for a tower, the camera path is rather circular. Straight paths are usually ruled out to avoid large pixel ground size variability and grazing point views.

But, it is less easy during the fieldwork to check that the overlap between two successive images is well respected. Indeed, for that type of object, the aspect is often repetitive and the vision could be deceived because of the round shape. Thus, to correctly acquire the images, we suggest placing marks at the places where the images must be taken. These places are separated by the baseline distance.

In this context, we present the formula to compute this base for different implementations in the field (part 3). Fieldworks are made on a quite small round tower (part 4). We then discuss the results (part 5). This work concerns the exterior and the interior of the building.

## 3. FORMULATION

### 3.1 Outdoor image capture

We propose a simple formula to compute the curvilinear baseline (figure 1). It is expressed in terms of the desired covering, the camera intrinsic parameters and the approximate cylinder radius.

* Corresponding author

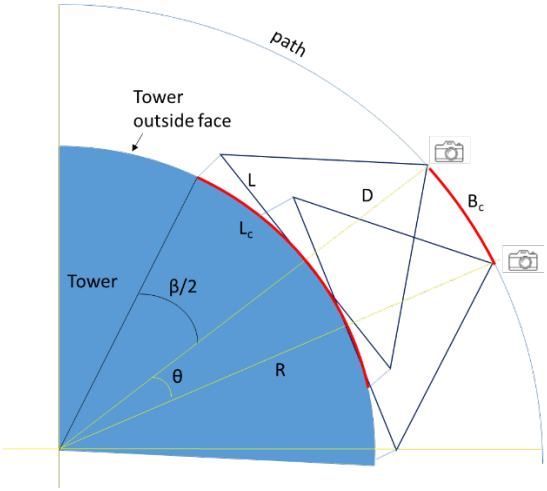


**Figure 1.** Top view of a round tower showing a circular camera path. $B_c$ is the curvilinear baseline. $L_c$ is the curvilinear length of the image footprint, and L the equivalent straight image footprint. D is the capture distance. R is the radius of the tower. θ is the angle from two successive camera positions.

The curvilinear length of the image footprint is computed with :

$$L_c = 2R \arcsin\left(\frac{L}{2R}\right) \tag{2}$$

where $R$ = tower radius
$L$ = tangential length of the image footprint

One can show that :

$$\theta = \frac{L_c(1-r_x)}{R} \tag{3}$$

where θ = angle between two successive camera positions measured from the centre of the cylinder
$r_x$ = longitudinal overlap rate between two images

Then, the curvilinear baseline is written :

$$B_c = (R + D)\theta \tag{4}$$

As it is not easy to measure the desired curvilinear baseline in situ, the value could be converted to the rope length of the arc :

$$B = 2(R + D)\sin\frac{\theta}{2} \tag{5}$$

With this baseline, we discuss three strategies of image acquisition: only one frame is captured from each station, a set of three frames (frontal image, looking upwards and looking downwards), or a set of five frames (frontal image, looking backwards and looking forwards, looking upwards and looking downwards). Let us note these strategies E1, E2 and E3 respectively.

Hence, adding more images increases the point visibility redundancy for both the SfM and point cloud densification steps, and it makes possible to cover the entire height of the wall. But these oblique views have grazing angles which also could degrade the point matching.

### 3.2 Indoor image capture

Concerning the interior surfaces, we also compare three strategies (figure 2). In the first one, the camera follows a circular path and looks towards the nearest wall. Let us note this strategy I1. In the second one, the camera follows a circular path but looks towards the furthest wall (strategy I2). For the last strategy, noted I3, the camera follows several perpendicular straight paths, leading to a larger number of acquired images.

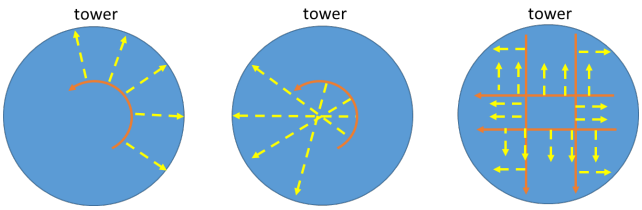


**Figure 2.** Three strategies of indoor capture, noted respectively I1, I2 and I3.

The formula of the baseline are different for each case. For the I1 strategy (figure 3), we have:

$$L_c = 2D \arcsin\left(\frac{L}{2D}\right) \tag{6}$$

$$B_c = (R - D)\theta \tag{7}$$

where θ = given in equation 3 where $L_c$ is from equation 6

$$B = 2(R - D)\sin\frac{\theta}{2} \tag{8}$$

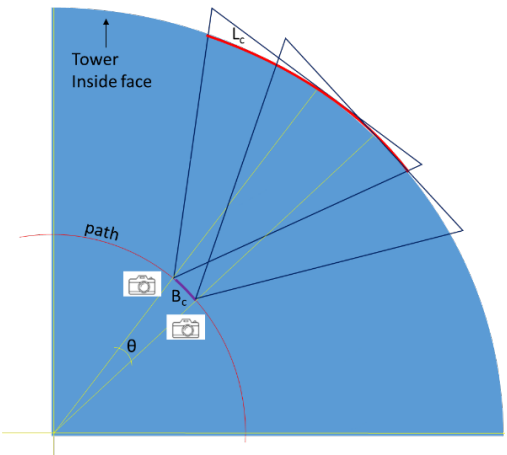


**Figure 3.** Top view showing the $B_c$-baseline inside the tower for I1 strategy.

For the I2 strategy (figure 4), we have:

$$B_c = (D - R)\theta \tag{9}$$

where θ = given in equation 3 where $L_c$ is from equation 6

$$B = 2(D - R)\sin\frac{\theta}{2} \tag{10}$$

For the I3 strategy, we recall that the base is given by:

$$B = (1 - r_x)L \tag{11}$$

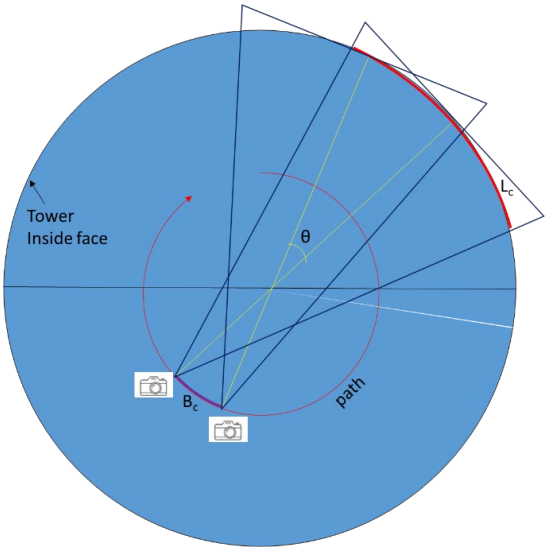


**Figure 4.** Top view showing the $B_c$-baseline inside the tower for I2 strategy.

Here, we take three frames per camera station (frontal image, looking upwards and looking downwards) to cover the entire height of the surface. Contrary to the outside case, there is no grazing angle effect.

## 4. DATA ACQUISITION AND PROCESSING

### 4.1 Studied object and fieldwork

The experiments are made on a rather small cylindrical object (figure 5). It allows us to test the different exposed strategies where the baseline is computed according to the above equations. The tower radius is around 3.6m inside and 3.8m outside. The height is ranging from 2.9m to 4.4m inside, and 5m outside. The faces are made of bricks. A wooden frame is visible inside.

**Figure 5.** Tested cylindrical tower : outside and inside.

We use the Canon EOS 850D with a 18mm focal length. The image size is 6000x4000 px. The pixel size is 3.7211µm. Ground control points (GCP) are surveyed using a total station Leica TS06+ whose angular accuracy is 0.3mgon and distance accuracy is 2mm+2ppm. GCP are recorded by free stationing and double measurements.

During the fieldwork, we decide to fix two parameters: (i) the overlap between successive frontal views is set to 70%, (ii) the overlap between a frontal image and the two associated oblique images is set to 50%. Here, we follow some guidelines found in (Caudal, 2021). Unless they were not formulated for a round object., this last parameter is not assessed here.

### 4.2 Outdoor acquisitions

We arbitrary fix the GSD to 1mm. According to equation 1, the shooting distance, D, is 5m. The base, B, is 7.44m (equation 5). This leads to 9 camera stations, so 9 images for E1 strategy (figure 6), 27 for E2 and 45 photos for E3 (figure 7). This data is summarized in the table 1. 15 GCP are measured around the tower at different heights (figure 8). Points 3, 4, 7, 9 and 12 are used as check points.

| Strategy | E1 | E2 | E3 |
|---|---|---|---|
| D (m) | 5 | 5 | 5 |
| GSD (mm) | 1 | 1 | 1 |
| θ (degree) | 33 | 33 | 33 |
| B (m) | 5.0 | 5.0 | 5.0 |
| Number of camera stations | 9 | 9 | 9 |
| Number of images | 9 | 27 | 45 |

**Table 1.** Acquisition parameters in the outdoor context.

**Figure 6.** The nine photos for the E1 strategy.

**Figure 7.** Three photos for the E2 strategy (frontal, downwards, upwards), plus the two additional photos for the E3 strategy (backwards, forwards).

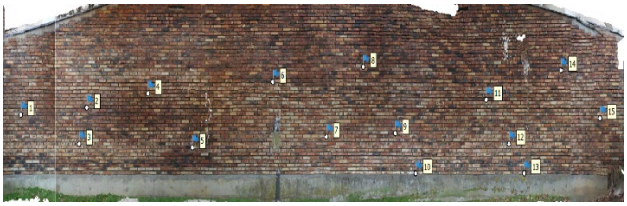

**Figure 8.** GCP locations (blue flags) on an unfolded image of the rounded tower (outside).


https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1473-2023  

### 4.3 Indoor acquisitions

Here the GSD cannot be the same for the three strategies. It will be lesser for I3 and I1 than for I2. The shooting distance is constrained by the radius of the tower. The table 2 summarizes the distance and GSD that are implemented in our case.

| Strategy | I1 | I2 | I3 |
|---|---|---|---|
| D (m) | 2.6 | 4.6 | 3 |
| GSD (mm) | 0.5 | 1.0 | 0.6 |
| θ (degree) | 17 | 29 | x |
| B (m) | 0.3 | 0.5 | 1.1 |
| Number of camera stations | 17 | 13 | 20 |
| Number of images | 51 | 39 | 60 |

**Table 2.** Acquisition parameters in the indoor context.

A sample of images are displayed on figures 9, 10 and 11. As previously, many GCP (13 points) are surveyed along the surface with different heights (figure 12). Points 8, 10, 12, 2 and 3 are used as check points.

**Figure 9.** Some images acquired for the I1 strategy.

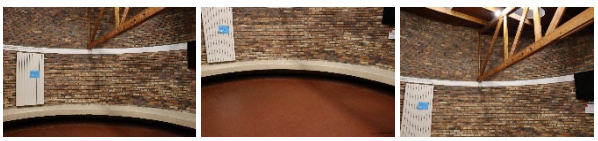

**Figure 10.** Some images acquired for the I2 strategy.

**Figure 11.** Some images acquired for the I3 strategy.

**Figure 12.** GCP locations (blue flags) on an unfolded image of the rounded tower (inside).

### 4.4 Data processing

Topographic data are processed with the v17 Sogelink Covadis software with a millimetre accuracy in a local coordinates system.

Photogrammetric processing is done with the v2.0.1 Agisoft Metashape software using the following parameters: high image alignment, ultra-high (outside) or high (inside) point cloud densification with a moderate filtering, high quality of the meshing.

## 5. RESULT ANALYSIS

### 5.1 Comparison of the strategies outside the tower

Table 3 summarizes the results after the image alignment and the dense point cloud step. As expected, more tie points are generated using more photos and the bottom and the top of the object are better recovered with E2 and E3 (figure 14).

| **Strategy** | **E1** | **E2** | **E3** |
|---|---|---|---|
| **Alignment** | | | |
| Number of tie points | 13576 | **47920** | 45671 |
| Average tie point multiplicity | 2.0 | 2.6 | **2.8** |
| Coverage area (m²) | 2.4 | **29.1** | 25.2 |
| Reprojection error (px) | **0.6** | 2.2 | 0.9 |
| Control points RMSE (mm) | **5.9** | 12.1 | 6.7 |
| Check points RMSE (mm) | **4.6** | 6.8 | 5.1 |
| Control points error (px) | **0.8** | 2.0 | 1.8 |
| Check points error (px) | **0.5** | 2.3 | 1.6 |
| **Dense point cloud** | | | |
| Number of points (millions) | **77.4** | 45.3 | 43.3 |
| Resolution (mm/pix) | 5.3 | **4.9** | 5.2 |
| Point density (pts/cm²) | 3.6 | **4.2** | 3.7 |

**Table 3.** Results after the image alignment and dense point cloud computation (outside). Bold numbers are the most favourable.

However, the alignment for these two last strategies was not easy. We were confronted to many image misalignments or large errors of camera position estimation when using all images at once. It was necessary to proceed sequentially for this step. Then, as a result, the dense point cloud presents noisy points and large holes (figures 13 and 14). The metre accuracies are also degraded compared to E1. And the E2 and E3 orthomosaics contain geometric deformations (figure 15).

We think that these phenomena are due of the grazing angles between the object surface and the oblique images. It turns to be a source of error in the matching process. Furthermore, it may be also explained by a too small image overlap (70%).

**Figure 13.** Phenomenon of noisy points with E3 that does not exist with E1.

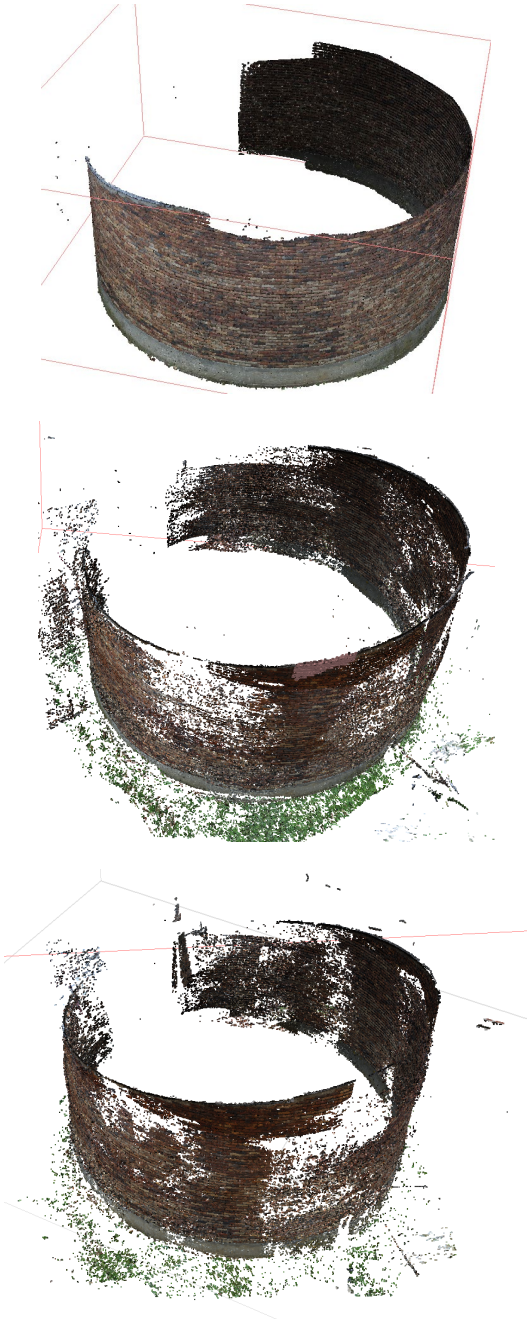

**Figure 14.** Dense point clouds from E1, E2 and E3.

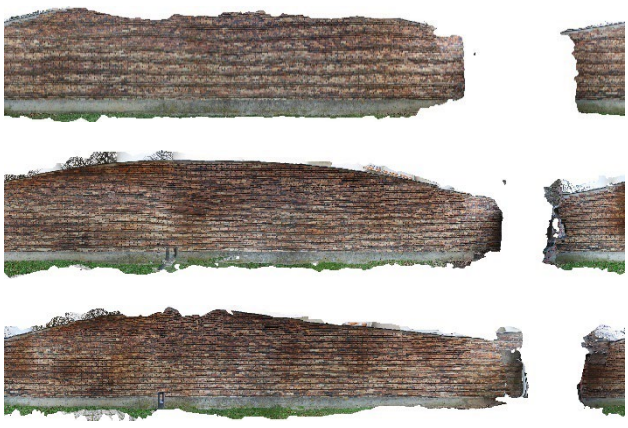

**Figure15.** Orthomosaics from E1, E2 and E3.

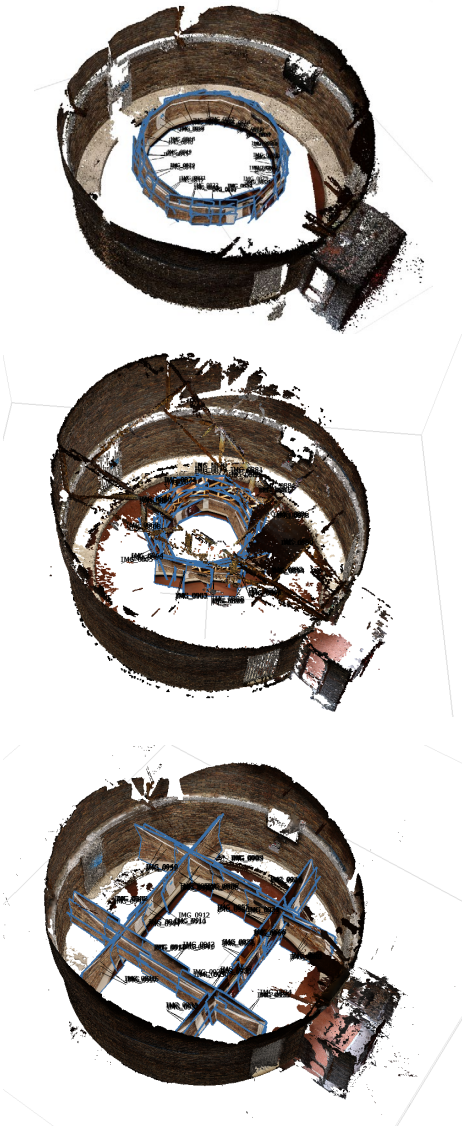

**Figure 16.** Views of the dense point clouds and camera positions for I1, I2 and I3.

**Figure 17.** Orthomosaics from I1, I2 and I3.

### 5.2 Comparison of the strategies inside the tower

According to the table 4, the I3 strategy, including more images, leads to the denser point cloud (figure 16) but turns to have the highest metric errors. The I1 strategy makes it possible to have a dense point cloud and higher meter accuracy. This can be explained as the GSD is better. However, the covered area is the highest with I2.

| **Strategy** | **I1** | **I2** | **I3** |
|---|---|---|---|
| **Alignment** | | | |
| Number of tie points | 63147 | 33346 | **81470** |
| Average tie point multiplicity | 3.3 | **4.5** | 3.0 |
| Coverage area (m²) | 16.0 | **33.8** | 15.8 |
| Reprojection error (px) | **0.8** | 0.9 | 1.0 |
| Control points RMSE (mm) | 6.0 | **4.0** | 39.2 |
| Check points RMSE (mm) | **6.1** | 6.7 | 22.6 |
| Control points error (px) | **0.6** | 0.8 | 1.3 |
| Check points error (px) | **0.6** | 0.8 | 1.7 |
| **Dense point cloud** | | | |
| Number of points (millions) | **119.6** | 22.3 | 88.6 |
| Resolution (mm/pix) | **1.0** | 5.1 | **1.0** |
| Point density (pts/cm²) | 1.0 | 3.9 | **95.1** |

**Table 4.** Results after the image alignment and dense point cloud computation (inside). Bold numbers are the most favorable.

Unless there are less captured images with I2, the orthomosaic clearly shows less occlusions (figure 17). We also note obvious deformations in the I3 orthomosaic.

Then, for an indoor acquisition, when possible, the I2 strategy could be recommended. This confirms some art rules as in (Agisoft, 2022).

## 6. CONCLUSIONS

In the framework of close-range photogrammetry, we assessed different strategies for the image capture of a round tower and indicated the computation of the adapted baseline given a capture distance, a camera and an overlap rate.
Outside, among the three tested strategies, we confirm that a single photo taken towards the centre of the tower should be privileged, but with a higher overlap rate than 70% used in this work. Adding oblique images has not proven to be beneficial.
Inside, among the three presented strategies, the circular path should be preferred. Then, shooting towards the nearest or the farthest sides will depend on the size of the tower. If both are possible, we obtained a better coverage using the second strategy.

## ACKNOWLEDGEMENTS

The author greatly thanks the four ESGT students: G. Batifol, T. Guivarch, T. Le Goff and C. Senra who participated to this project. They have acquired all the data used in this work and made the data processing.

## REFERENCES

Agisoft LCC, 2022. *Agisoft Metashape User Manual Professional Edition*, Version 1.8.

Ahmadabadian, A., H., Robson, S., Boehm, J., Shortis, M., Wenzel, K., Fritsch, D., 2013. A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs, *ISPRS J. of Photogramm. Rem. Sens.,* 78, 157–167.

Boehler, W., Marbs, A., 2004. 3D scanning and photogrammetry for heritage recording: A comparison. *Proc. 12th Int. Conf. on Geoinformatics*. Sweden, 291–298.

Caudal, P., 2021. *Apport de la photogrammétrie et de l'intelligence artificielle à la détection des zones amiantées sur les fronts rocheux*. PhD University of Le Mans, France.

García-Gago, J., González-Aguilera, D., Gómez-Lahoz, J., San José-Alonso, J. I., 2014. A Photogrammetric and Computer Vision-Based Approach for Automated 3D Architectural Modeling and Its Typological Analysis. *Rem. Sens.*, 6(6), 5671-5691.

Grussenmeyer, P., Landes, T., Voegtle, T., Ringle, K., 2008. Comparison methods of terrestrial laser scanning, photogrammetry and tacheometry data for recording of cultural heritage buildings. *Int. Archives of Photogramm., Rem. Sens. Spatial Inf. Sci.* Vol. XXXVII. Part B5. Beijing, 213–218.

Kraus, K., 2007. *Photogrammetry, Geometry from Images and Laser Scans*. De Gruyter, Berlin, 2nd edition.

Marelli, D., Bianco, S., Ciocca, G., 2020. IVL-SYNTHSFM-v2: A synthetic dataset with exact ground truth for the evaluation of 3D reconstruction pipelines. *Data in brief*, 29, 105041.

Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V., Chamzas, C., 2007. Methods for 3D digitization of Cultural Heritage. *J. of Cultural Heritage*, 8, 93–98.

Tuttas, S., Braun, A., Borrmann, A., Stilla, U., 2016. Evaluation of acquisition strategies for image-based construction site monitoring. *Int. Arch. of the Photogramm., Rem. Sens. Spatial Inf. Sci.*, XLI-B5, 733–740.

Waldhäusl, P., Ogleby, C., 1994. 3 x 3 Rules for simple photogrammetric documentation of architecture. *International Arch. of Photogramm. Rem. Sens.*, Vol. XXX, Part5, 426–429.

Wenzel, K., Rothermel, M., Fritsch, D., and Haala, N., 2013. Image acquisition and model selection for multi-view stereo. *Int. Arch. Photogramm. Rem. Sens. Spatial Inf. Sci*., XL-5/W1, 251–258.