High-Speed 3D-Documentation of Schönbrunn Palace

Pushing technological boarders in completeness, resolution, and accuracy

Peter DORNINGER, 4D-IT GmbH, Austria
Nikolaus STUDNICKA, RIEGL Laser Measurement Systems GmbH, Austria

Abstract: Detailed, complete and accurate 3D-documentation is of major importance for adequate management of cultural heritage sites. Respective models can be used for many applications such as mapping, status documentation before and after restoration or to generate digital twins based on Building Information Modelling (BIM) technology. A variety of sensors and techniques emerged in the past two decades, ranging from automated, low budget, out-of-the-box solutions using photos from mobiles phones up to cutting edge technologies enabling survey-grade results at millimetre scale and accuracy for entire buildings or sites. In this paper, an approach based on fast and efficient terrestrial laser scanning (TLS) for façade documentation is presented. High-up areas and the roof landscape are captured using a UAV-driven medium-format camera. By means of this technology, the entire main building of *Schönbrunn Palace* (~180 × 60m) may be documented within one working day. Sensor technology (TLS: *RIEGL* VZ400i, UAV-camera: *Phase One* iXM-100) and data acquisition strategies as well as a quick and efficient visualization approach are described. In addition, different methods for thinning point-clouds are presented and evaluated. The latter are well suited for high-detailed triangle-mesh generation. Finally, the potential of the LiDAR-sensors integrated into the latest generation of Apple's iPhone is evaluated for complementary 3D-data acquisition and modelling.

Keywords: *Terrestrial Laserscanning—UAV—High-Resolution Camera—Automation—3D Mesh*

CHNT Reference: Dorninger, P. and Studnicka, N. (2022). 'High-Speed 3D-Documentation of Schönbrunn Palace', in Börner, W., Rohland, H., Kral-Börner, C. and Karner, L. (eds.) *Proceedings of the 25th International Conference on Cultural Heritage and New Technologies, held online, November 2020.* Heidelberg: Propylaeum.
doi:10.11588/propylaeum.1045.c14493

Introduction

Schönbrunn Palace was the main summer residence of the Austrian emperor and is located in Vienna, Austria. Since 1996, it is listed as *UNESCO World Heritage Site* and, in addition, being one of the most important and most visited touristic hot-spots of Vienna, special requirements on continuous monitoring and documentation are given. Especially ongoing restoration and construction work require detailed, complete and accurate 3D-documentation.

Since 2000, the *Schloß Schönbrunn Kultur- und Betriebsges. m.b.H* who is operating and managing the heritage site established a continuous procedure for 3D-documentation of all objects being restored. Terrestrial (TLS) and airborne laserscanning (ALS) as well as image based structure from motion (SFM) is commonly used to document the status before the restoration as foundation for decision making, and the status after the restoration for long-time monitoring as well as for the case that something is destroyed by an unforeseen event (e.g., Dorninger et al., 2013).

In this contribution, a highly automated approach is described, based on cutting edge technology to enable the 3D-documentation including high-resolution image acquisition of the exterior of the main building of *Schönbrunn Palace* (size: ~180 × 60 m / 6,800 m²; façade length: ~550 m) within one day. Modelling, visualization and analysis of objects in the fields of cultural heritage and infrastructure monitoring will be discussed. Special focus is drawn on hybrid acquisition techniques based on TLS providing a highly accurate reference and integrating high-accuracy, high-resolution image data from a UAV-driven *Phase One* iXM-100 camera to capture high-up areas and the entire roof-landscape of the building. The presented 3D-modelling approach is based on sophisticated 3D-meshing techniques. The highly automated workflow provides models perfectly suitable for Building Information Modelling (BIM) in both, accuracy and resolution. Finally, the potential of the LiDAR sensors integrated into the latest generation of *Apple*'s iPhones is evaluated for complementary data acquisition.

Related Work

Image based solutions for 3D-reconstruction have become very popular and hence the application of respective software tools is a widespread technology. Remondino et al. (2017) compare three commonly used image processing tools (*Agisoft Metashape*, *Pix4D*, *Capture Reality*) with respect to the achievable accuracy of all necessary processing steps including bundle adjustment (BA) and SfM and they define several measures to evaluate accuracy and reliability of such methods. Especially for UAV-based applications, so called videogrammetry approaches are emerging. To simplify the data-acquisition process (i.e. selecting the view-points), entire video streams are captured and subsequently, the individual frames are processed automatically to generate 3D-point clouds. Torresani & Remondino (2019) analyse the differences between videogrammetry and "conventional" photogrammetric approaches and they describe different methods to process the frames in order to increase the achievable quality of the resulting models. They conclude their work with "*3D documentation of heritage scenarios with high-end high-resolution cameras will be never surpass*"; especially if high-resolution, high-quality, and high accuracy are of major demand. Hence, for small to medium sized objects (e.g., sculptures, fountains, ...) or for parts of buildings (some 100 m of facades, distinct parts of a roof, …) or if the visual aspect is of major interest, such approaches may be the best solution considering effort (time and financial) and achievable result. However, if aiming at high-resolution and especially high accuracy, the approach proposed in the following will be superior. Incredible progress in sensor technology development enables the acquisition of some 400 scans per day by a single operator including on-site registration of the collected data. The *RIEGL* VZ-400i laser scanner measures with millimetre resolution and precision and integrates GNSS and IMU sensors and a powerful processor for this task. Multi Station Adjustment, a fine-registration post-processing step, copes with reference points to increase the absolute accuracy and to enable georeferencing of a project (Ullrich, 2017).

Data Acquisition and Processing

The data acquisition is based on the following three steps:

1. Geodetic network measurement providing precise positions of 14 retroreflecting marker points defining a local reference frame in relation to the Austrian reference frame.
2. Terrestrial Laser Scanning (TLS) using a high-speed and high-resolution *RIEGL VZ-400i* for capturing the facades in 3D and colour.

3. UAV-based image acquisition using a high-resolution *Phase One iXM-100* medium format camera for capturing high-up areas and the roof-landscape in 3D and colour.

For this project, the work was carried out on 3 days in March (UAV-based image acquisition) and April 2020 (TLS). The entire TLS data acquisition was carried out by a single operator within 6.5 hours. For the UAV operation, it is advisable to have two persons, one controlling the UAV and one checking the camera settings and viewing positions. The geodetic network can be realized by one operator, however, two are advisable for efficiency reasons (i.e. switching totalstation and mirrors between the tripods). Therefore, considering a team of 4–5 people, the complete data capturing would be possible within one working day as the different processes do not handicap each other. With a smaller team (e.g., 2 persons), two days are realistic. Figure 1 shows the TLS with camera (Nikon D-850) and GNSS, and the UAV equipped with the *Phase One iXM-100* camera.

a *b*

Fig. 1. a) TLS RIEGL VZ-400i placed on the main staircase of the northern courtyard of Schönbrunn palace. The TLS and the on-top mounted camera have been used to capture a high resolution, colour coded point-cloud of the facades; b) DJI Matrice 600 pro with Phase One iXM-100 medium format camera as used for capturing high-resolution stereo-photos of the roof area. © Authors.

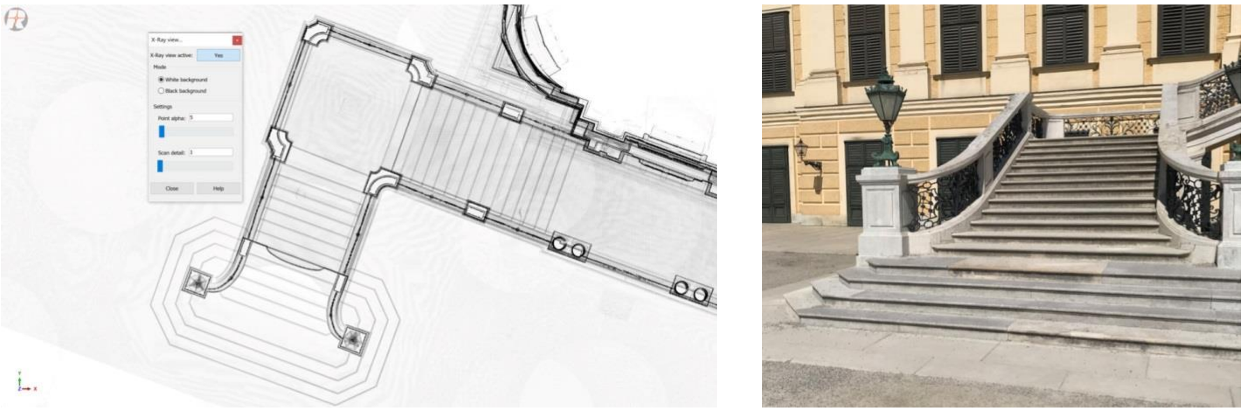
Terrestrial Laserscanning

The typical workflow of the eye-safe *RIEGL* VZ-400i laser scanner used here is to acquire one so-called panoramic scan after another. The scanning parameters have been selected to maximize the recording efficiency. A so-called “Panorama40” scan (360° × 100° field of view, 40 milli-degrees resolution) requires 45 seconds scanning time. An average of 22.5 million point measurements per scan is performed. The spatial resolution of the measuring points at a distance of 10 meters is 7 mm. Five 45-megapixel photos are taken per scan. The time required for the 292 scanning positions was 6.5 hours for a single operator. Table 1 lists the specification of the scanner.

Table 1. Technical specification of the TLS and camera sensor configuration used for the façade modelling.

Laser scanner	*RIEGL* VZ-400i
Field of view of the laser scanner	100° vertical x 360° horizontal
Angular resolution	0.040° (7 mm @ 10 m distance)
Precision of range measurement	3 mm
Attached photo camera	Nikon D-850 (14 mm lens), 45 MPix / image

Already within the laser scanner it is possible to combine the scan data of one scan after the other, thus to “register” the scan positions. This process is multi-stage, using a built-in GNSS receiver, built-in IMU (inertial measurement unit), and the subsequent *Fourier Transformation* and *ICP* (iterative closest point) algorithms (Ullrich, 2017). The post-processing of the scan data includes filtering the scans (e.g., eliminating so-called “ghost-points”), the fine block adjustment of all scan positions (“Multistation Adjustment”) by the use of the geo-referenced markers, colorize the scans from the photos and homogenization of the data (using Octree Extractor). A detail of the merged scandata of the southern staircase is shown in Figure 2.



X-Ray view active: Yes
● White background
○ Black background
Point alpha: 5
Scan detail: 1

a *b*

Fig. 2. a) X-Ray view of the TLS point-cloud highlighting vertical structures of a part of the southern staircase of Schönbrunn palace b) photograph of the same area of the staircase. © Authors.

UAV-based High-Resolution Camera

A variety of UAV-based data acquisition systems do exist ranging from “low-budget” systems using consumer camera-drones (e.g., Sun and Zhang, 2019) to heavy payload drones with calibrated, high-resolution cameras and direct georeferencing sensors. The specification of the *Phase One iXM-100* medium format camera used in this project is given in table 2.

Table 2. Technical specification of the Phase One medium format camera used on a UAV to capture the roof structure and texture.

Camera	Phase One iXM-100
Lens	RSM 35mm f/5.6 (63 × 49,4° opening angle)
Image resolution	100 MPix (11,664 × 8,750 Pixel)
Effective sensor size	43.9 × 32.9 mm

This high-end camera is superior to many competitors in several characteristics. The large sensor enables very high radiometric depths. The optimized RSM lenses reduce the number of optical elements to a minimum increasing the amount of light passing through as well as the effective resolution having positive effect on minimizing so called lens errors such as distortion or vignetting. At a mean flight height of 35 m above the roof landscape, an image resolution of 3.7 mm per pixel could be achieved. The entire roof was captured with a vertical camera position using waypoint flight. In addition, 45° tilted viewing direction towards the façade was used to complete the area along the balustrade. Due to the lighting conditions, it was necessary to plan the flights w.r.t. the position of the sun to prevent hard shadows. Figure 3 shows the potential of the sensor. The low sun elevation in March causes very dark areas at roofs facing towards southern direction while some parts of the image are still very high-lighted (e.g., sculptures). Left, an original image is shown while the centre image shows a post processed version where dark areas have been automatically highlighted while also increasing the colour depth in highlighted areas.

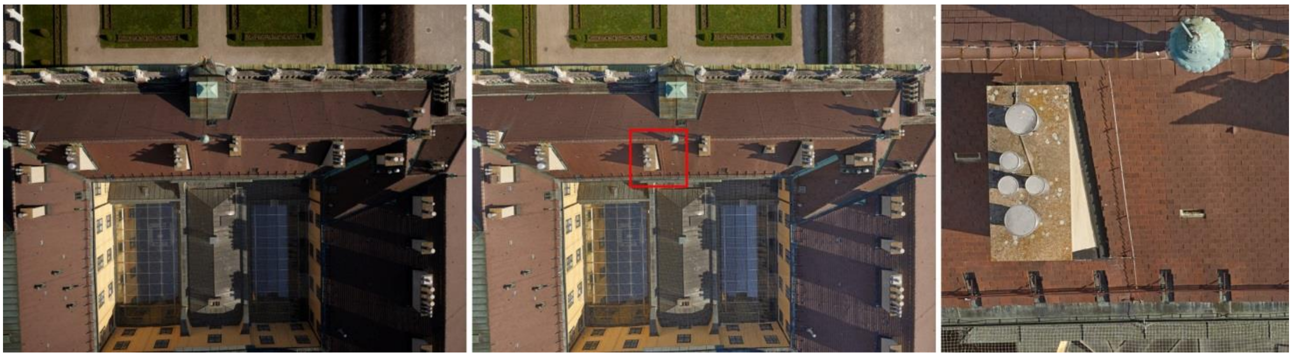


Fig. 3. Original 100 MPix image with dark, shadowed areas (left); post-processed version after applying radiometric correction and high-dynamic-range improvement (center); the bounding box of the detail (right) is highlighted as red rectangle in the center image. © Authors.

Georeferencing

Point-cloud based approaches as well as surface modelling (i.e. triangulation meshes) have been evaluated for representation, documentation and analysis of the data. As described, the TLS data has been referenced using a survey-grade network of marker-points. To reduce the amount of data, a homogeneously distributed point cloud with 10 mm resolution has been derived using a voxel based averaging approach. The UAV-image data has been processed using bundle-block adjustment and structure from motion (*Agisoft Metashape*). The resulting point-cloud has been fine referenced w.r.t. the TLS voxel-point cloud. For first visualization models, the coloured voxel-point-cloud of the façade area has been merged with a textured high-resolution mesh of the high-up areas and the roof landscape. To derive the mesh from the SfM-point cloud, an advanced implementation of the Poisson triangulation has been used as described by Nothegger (2011). Figure 4 shows the result of this visualization approach.

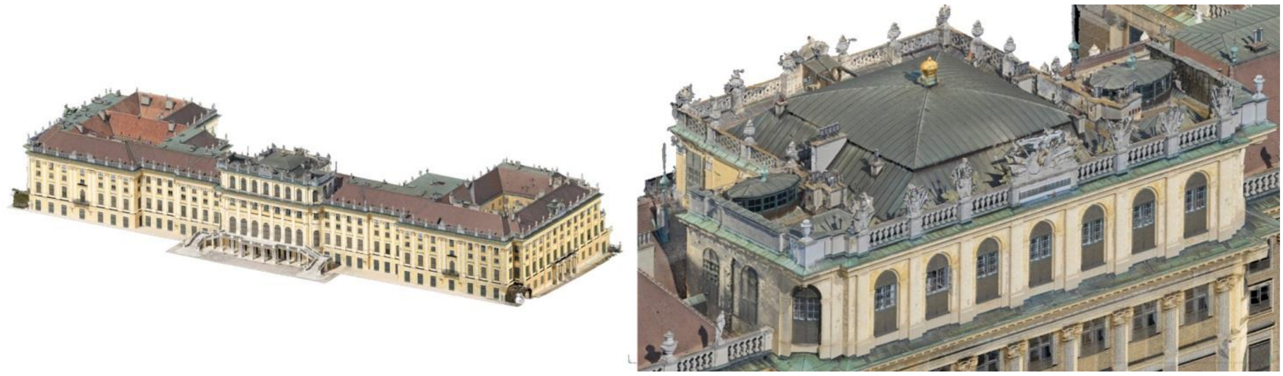


Fig. 4. Textured 3D-model of Schönbrunn Palace combining the TLS point-cloud at the façade area and a textured 3D-triangulation model of the roof structure. © Authors.

The following Table 3 summarizes the achieved accuracy of the georeferencing approach.

Table 3. Accuracy of co-georeferencing TLS and UAV data in comparison to a survey-grade control-point network.

Processing step	**Accuracy**
Multistation adjustment of 292 scanning positions using real-time GNSS/IMU position and orientation information as input	inner accuracy: ±1 mm
14 marker-points of a geodetic network defining the reference frame	inner accuracy: ±6 mm
Bundle-block adjustment of 458 cameras and 57 ground control points measured by GNSS or totalstation	inner accuracy: ±10 mm
Fine registration of SFM-point-cloud to TLS-point-cloud using identical, planar areas (facade, ground, roof)	absolute accuracy: ±10 mm

Point-cloud vs. Triangulation-Mesh

TLS and SfM approaches basically generate 3D-point-clouds. While TLS directly measures the 3D-points at the object's surface, SfM applies the stereoscopic measurement principle to derive 3D-surface-points. Dependent on the measurement instrument, the object's surface and several further parameters (e.g., texture, material, etc.), the individual points are not measured at their exact position. These randomly distributed errors are often referred to as measurement noise. In order to minimize this effect on the one hand and to reduce the amount of data on the other hand, averaging approaches are commonly applied (e.g., voxel-based averaging as described above). For numerous applications including visualization, distance measurement or mapping, such point-clouds can be applied "as they are" (see façade representation in Figure 4). Especially for real-time representation using web-streaming technology, point-clouds are well suited (e.g., Schütz et al., 2020). However, for several applications, 3D-surface models (i.e. mesh / triangulation models) are superior. They implicitly represent the surface topology, i.e., connected components and volumes can be determined, and volume differences (e.g., change detection), as well as sections can be computed easily. Furthermore, adaptive thinning (i.e. reducing the amount of data at smooth areas) can be applied to reduce the amount of data while preserving details at highly structured areas. To support the interpretability and for photorealistic appearance, meshes enable texture mapping with high-resolution images. The achievable quality of the mesh is dependent on the point-cloud. Figure 5 shows a Poisson triangulation of the voxel-based point-cloud (upper) and after preprocessing the data using the approach described by Nothegger (2011) (lower). It can clearly be seen, that the latter approach provides a superior result on smooth surfaces while preserving details at sharp edges or small structures.

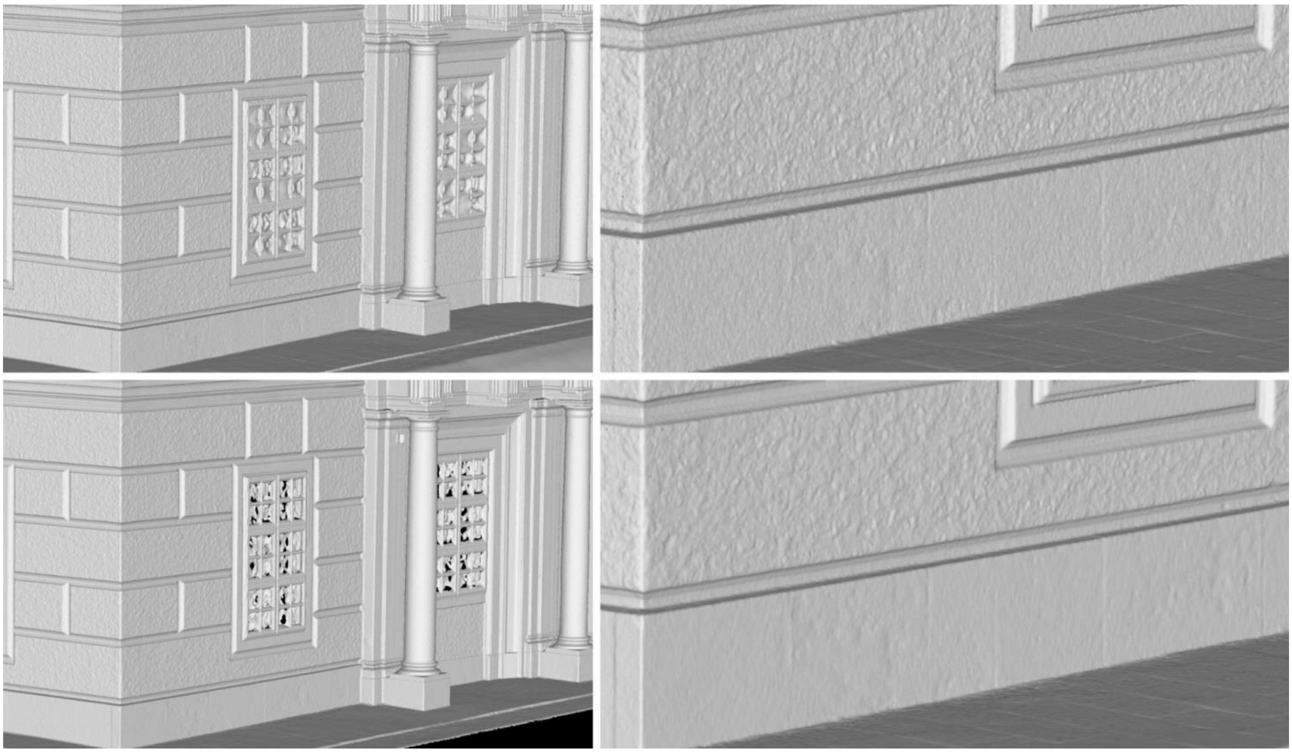


Fig. 5. Upper: Triangulation mesh of voxel-based point cloud with 1 cm resolution; Lower: Mesh derived from pre-processed point cloud. While the voxel-based approach is a simple average using a pre-defined voxel structure, the other approach applies a more sophisticated thinning and smoothing resulting in a better elimination of measurement noise while preserving structural detail. © Authors.

Fig. 6. Result of the complementary data acquisition using the iPhone-LiDAR sensor. f.l.t.r.: Original picture from iPhone; triangulation mesh from iPhone LiDAR sensor; textured 3D-mesh. © Authors.

Complementary Data Acquisition

The most recent generation of *Apple*'s iPhone is equipped with a LiDAR sensor. In combination with the high-quality camera of the mobile device, comparably high-resolution and textured 3D models can be derived automatically. For huge objects (e.g., entire buildings), the accuracy of this technology may not be sufficient. However, this kind of devices may enable in the near future for complementary data acquisition of parts of assets and an automated integration into available, highly accurate 3D-models. Evaluating the richness in detail in combination with the photo-realistic texture of the achievable models (Fig. 6), makes obvious, that, taking benefit of the automated workflow, inexperienced users might be able to generate 3D-models. By means of respective algorithms for automated registration and integration into available 3D-models (e.g., Multi-Station-Adjustment – Principal Component Analyses), this technology is very likely to integrate 3D-measurement technology into the daily work of people responsible for cultural heritage sites.

Discussion and Outlook

The described approach demonstrates that the integration of cutting-edge, high-resolution terrestrial and airborne data acquisition enables a complete documentation of huge and high objects in a very fast and hence efficient manner. The main building of *Schönbrunn Palace* can be captured in 3D within 1 day at a resolution and absolute accuracy of 10 mm and with 3.6 mm texture. We are aware, that compared to other "low-budget" approaches, the demand for the equipment is comparably high. However, considering restricted time budget (e.g., closing areas for tourists, weather conditions, legal restrictions, …), and due to the fact that field work is typically expensive, the high costs of the equipment can be overcome quickly, especially if high accuracy is required. Due to its complete coverage, the data is well suited for automated processing based on artificial intelligence and machine learning to derive data for Building Information Modeling. In addition, different approaches for point-cloud thinning as preprocessing for subsequent mesh-triangulation have been analyzed. The example given shows, that noise reduction is possible while preserving detail if an appropriate approach is applied. Finally, a low budget solution for 3D-data acquisition and modeling, based on the most recent iPhone-LiDAR-technology was presented. By means of respective automated geo-referencing algorithms, this will enable soon to complement existing, survey-grade 3D-models by inexperienced users

Acknowledgments

The project was realized on behalf of *Schloß Schönbrunn Kultur- und BetriebsgesmbH*. This work was supported by *Phase One Industrial*. The UAV flight was realized by *Gecko4K OG*.

Conflict of Interests Disclosure

The authors declared no conflict of interests.

Author Contributions

The authors contributed equally to all works related with this paper.

References

Dorninger, P., Nothegger, C., and Rasztovits, S. (2013). ‘Efficient 3D Documentation of Neptune Fountain in the Park of Schönbrunn Palace at Millimeter Scale’. Talk: XXIV International CIPA Symposium 2013, Strasbourg, in “ISPRS-Annals Vol II-5/W1 (2013)”, pp. 103–108.

Nothegger, C. (2011) ‘Improving Completeness of Geometric Models from Terrestrial Laser Scanning Data’, in: Proceedings of the 23rd CIPA Symposium, Prague, Czech Republic.

Remondino, F., Nocerino, E., Toschi, I., and Menna, F. (2017). ‘A critical review of automated photogrammetric processing of large datasets’, ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-2/W5. 591–599. doi:10.5194/isprs-archives-XLII-2-W5-591-2017.

Schütz, M., Ohrhallinger, S., and Wimmer, M. (2020). ‘Fast Out-of-Core Octree Generation for Massive Point Clouds’, Computer Graphics Forum, 39(7), pp. 1–13, November 2020.

Sun, Z. and Zhang, Y. (2019). ‘Accuracy Evaluation of Videogrammetry Using a Low-Cost Spherical Camera for Narrow Architectural Heritage: An Observational Study with Variable Baselines and Blur Filters’, Sensors Vol. 19(3). doi:10.3390/s19030496.

Torresani, A. and Remondino, F. (2019). ‘Videogrammetry vs photogrammetry for heritage 3D reconstruction’, ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII-2/W15. **, 1157–1162. doi:10.5194/isprs-archives-XLII-2-W15-1157-2019.

Ullrich, A. and Fürst, Ch. (2017). ‘Vollautomatischer Ansatz für die Onboard-Datenregistrierung im terrestrischen Laserscanning’, Beiträge zum 165. DVW-Seminar am 11. und 12. Dezember 2017 in Fulda.