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

CULTURAL HERITAGE DIGITAL PRESERVATION THROUGH AI-DRIVEN ROBOTICS

G. Marchello[2], R. Giovanelli[1-3], E. Fontana[2], F. Cannella [2], A. Traviglia [1]*

[1] Istituto Italiano di Tecnologia, Center for Cultural Heritage Technology, 30172 Venice, Italy – (arianna.traviglia, riccardo.giovanelli)@iit.it

[2] Istituto Italiano di Tecnologia, Center for Convergent Technologies - Industrial Robotics Facility, 16163 Genoa, Italy – (gabriele.marchello, eleonora.fontana, ferdinando.cannella)@iit.it

[3] Ca' Foscari University of Venice, DSU, 3246 Venice, Italy – (riccardo.giovanelli@unive.it)

KEY WORDS: Digital Twins, Robotics, Computer Vision, Structure from Motion, Artificial Intelligence, Cultural Heritage

ABSTRACT:

This paper introduces a novel methodology developed for creating 3D models of archaeological artifacts that reduces the time and effort required by operators. The approach uses a simple vision system mounted on a robotic arm that follows a predetermined path around the object to be reconstructed. The robotic system captures different viewing angles of the object and assigns 3D coordinates corresponding to the robot's pose, allowing it to adjust the trajectory to accommodate objects of various shapes and sizes. The angular displacement between consecutive acquisitions can also be fine-tuned based on the desired final resolution. This flexible approach is suitable for different object sizes, textures, and levels of detail, making it ideal for both large volumes with low detail and small volumes with high detail. The recorded images and assigned coordinates are fed into a constrained implementation of the structure-from-motion (SfM) algorithm, which uses the scale-invariant features transform (SIFT) method to detect key points in each image. By utilising a priori knowledge of the coordinates and SIFT algorithm, low processing time can be ensured while maintaining high accuracy in the final reconstruction.
The use of a robotic system to acquire images at a pre-defined pace ensures high repeatability and consistency across different 3D reconstructions, eliminating operator errors in the workflow. This approach not only allows for comparisons between similar objects but also provides the ability to track structural changes of the same object over time.
Overall, the proposed methodology provides a significant improvement over current photogrammetry techniques by reducing the time and effort required to create 3D models while maintaining a high level of accuracy and repeatability.

1. INTRODUCTION

The use of 3D measurements and digital reconstruction in the domain of cultural heritage has become increasingly important in recent years due to technological advancements and greater accessibility to technologies that produce satisfactory results (Ćosović and Maksimović, 2022). Digitally rendered artifacts play a significant role in safeguarding material culture, including small objects, grand architecture, and entire cultural heritage sites. This is evident in the increasing efforts to systematically digitise global cultural heritage. For instance, the European Commission has recently introduced the common European data space for cultural heritage, a "new flagship initiative" funded under the Digital Europe Programme (DIGITAL) aimed at accelerating the digital transformation of Europe's cultural sector and promoting the creation and reuse of content in the cultural and creative industries (Europeana Foundation, 2023). This initiative highlights the significance of generating and sharing high-quality digital data from cultural heritage entities in a collaborative and accessible way. It is funded by the Digital Europe Programme (DIGITAL) of the European Union.

Producing digital twins of cultural heritage entities, which are virtual counterparts of physical products, assets, or systems that reflect the elements and dynamics of the way the complex systems run and evolve over time (Ćosović and Maksimović, 2022), is regarded as pressing and inevitable for purposes such as preservation, documentation, research, and public engagement. With a vast amount of cultural heritage around the world, there is an urgent need to thoroughly digitise these entities, which necessitates developing methods to reduce the time required to record, process, reconstruct, and deliver accurate 3D reproductions. It, in turn, demands automation in every phase of the 3D modeling pipeline (Remondino, 2011). The ultimate goal of 3D digitisation in cultural heritage is to create accurate, detailed, and accessible digital twins of physical objects, assets, and systems. These digital twins can be used for a variety of purposes, including preservation, documentation, research, and public engagement.

One of the significant benefits of 3D digitisation in cultural heritage is the ability to preserve and protect physical objects and sites. Digital twins can function as a backup in case of damage or destruction of the original object, and they can also be used to study and understand the object without risking damage. Additionally, 3D digitisation can provide valuable information for conservation and restoration efforts, enabling experts to analyse the structure and condition of the object and identify potential issues that may arise in the future.

Another essential benefit of 3D digitisation in cultural heritage is the ability to make these objects and sites accessible to a broader audience (Tsvetaeva, 2022). Digital twins can be shared online or through virtual reality experiences, enabling people from all over the world to experience and learn about these objects and

sites. This can be particularly valuable for objects and sites that are difficult to visit or are located in remote areas.

Furthermore, digital twins can play a crucial role in predictive maintenance of physical assets. By continuously monitoring the data from sensors installed on the asset, the digital twin can detect any anomalies and alert maintenance personnel before significant damage occurs (Luther et al., 2023). This can help prevent costly downtime and repairs.

Overall, digital twins have the potential to revolutionise cultural heritage management and become real "knowledge models" (Gabellone, 2022). As technology continues to evolve and become more sophisticated, we can expect to see even more innovative applications in the future.

In this article, after providing a comprehensive summary of current developments in both the acquisition and resolution of 3D models, we present a novel approach to produce precise and high-fidelity 3D models of archaeological artifacts by utilising 3D data acquisition techniques in combination with a computer vision system mounted on a robotic arm that follows a predetermined path. This ground-breaking technique not only minimises the need for manual labour but also ensures exceptional levels of accuracy and precision in the resulting models.

2. 3D DATA ACQUISITION

2.1 Acquisition methods

Various methods are used in the field of Cultural Heritage to acquire, elaborate, and store 3D models, including Structure-from-Motion (SfM), Structured Light Scanning, Laser Scanning, LiDAR, and others.

Structure-from-Motion (SfM) is a popular 3D reconstruction technique that recovers the 3D volume of an object from a series of images showing different views and recorded by one camera (Tomasi and Kanade, 1992). This methodology involves several steps, such as camera movement around the object, acquisition of multiple images, identification of features, matching, and assignment to a position in three-dimensional space. The processing time and the resolution of the reconstructed volume are proportional to the number of different views captured. The higher the number of captured profiles, the higher the level of detail and the processing time required to reconstruct the object in 3D. However, capturing a higher number of profiles also leads to longer digitisation time, which can be a constraint in reconstructing a large number of entities with the proper quality.

Structured Light Scanning is another 3D reconstruction technique that uses a projector and a camera to capture a series of patterns projected onto an object from different angles. The patterns create shadows on the object's surface, which are captured by the camera and used to reconstruct a 3D model. Structured Light Scanning has several advantages, including high accuracy, the ability to capture color information, and fast scanning speed. However, this method also has some drawbacks, including sensitivity to ambient light, limited range, and difficulty in capturing fine details (Rachakonda et al., 2019). Additionally, the equipment required for this technique can be expensive, and the setup process can be time-consuming.

Laser Scanning is a popular technique for 3D digital documentation and preservation of cultural heritage assets. The method involves a laser beam that is directed onto the object, which records the geometry and texture of the surface by measuring the time it takes for the laser to reflect back to the scanner. The resulting point cloud data can be used to create high-resolution 3D models of objects, buildings, and entire heritage sites, providing valuable information for research, conservation, and public engagement. Laser scanning can capture complex shapes and details that may be difficult to obtain with other techniques. However, laser scanning can be expensive, requires technical expertise to operate, and may not be suitable for objects that are sensitive to light or heat (Koch et al., 2017).

LiDAR (Light Detection and Ranging) is another remote sensing technology widely used in cultural heritage applications. LiDAR sends pulses of light to the surface of an object and measures the time taken for the reflected signal to return, allowing the creation of high-resolution 3D point clouds of the object. LiDAR can quickly capture large areas and is particularly useful for outdoor archaeological sites and large structures. It is also capable of capturing details of inaccessible areas, such as the interior of caves or the tops of buildings. However, LiDAR can be expensive and requires specialised equipment, making it less accessible than other 3D scanning methods. Additionally, its accuracy can be affected by atmospheric conditions and vegetation, which can result in incomplete or distorted data.

Apart from the aforementioned methods, other techniques used for 3D digitisation in cultural heritage include photogrammetry, multi-view stereo (MVS), and time-of-flight (ToF) cameras.

Photogrammetry involves capturing multiple photographs of an object from different angles and using software to reconstruct a 3D model.
MVS is a technique similar to SfM that involves capturing images of an object from multiple viewpoints and using computer algorithms to reconstruct a 3D model.

ToF cameras are sensors that emit infrared light and measure the time it takes for the light to reflect back, providing depth information that can be used to create 3D models.

Each of these techniques has its advantages and disadvantages, and the choice of technique will depend on factors such as the size and complexity of the object, the desired level of detail, and the available resources.

In heritage studies and archaeological practice, SfM is still considered the standard due to the increasing quality levels reached by photo cameras and the lower costs (Chandler and Buckley, 2016). Additionally, the superior quality of textures obtained through SfM technique, which is of utmost importance to archaeologists and heritage experts, make SfM an attractive option (Kaneda et al., 2022). However, SfM has its limitations, such as the slow processing when compared to other methodologies, and the potential for operator inaccuracy in the process of capturing the needed pictures.

The use of advanced technologies like AI-powered robotics can automate the implementation process of some of these 3D digitisation techniques. Robotics has become increasingly necessary in various industries due to the demand for efficiency, accuracy, and cost-effectiveness. The field of Cultural Heritage has similar needs. Automation can significantly enhance the implementation process of the data acquisition technique by reducing the manual labor involved in capturing and processing data, which can be time-consuming and prone to human error. Moreover, automation can enable the processing of large amounts of data in a fraction of the time it would take manually,

allowing for a quicker and more comprehensive analysis of cultural heritage assets.

Therefore, given the current state of the art in 3D modelling techniques and the need for faster and more accurate methods, the use of automation is an attractive option. By automating the acquisition techniques, we can minimise the limitations of manual effort and improve the quality and efficiency of 3D modelling processes.

To achieve this objective, we propose a workflow that integrates AI-powered robotics into the SfM technique. This workflow aims to automate the entire process of capturing images, identifying features, matching them, and assigning them to a position in three-dimensional space. By automating this process, we can achieve faster and more accurate results with minimal human intervention, thus increasing the productivity and efficiency of the 3D modelling process.

2.2 Resolution

Resolution is a crucial factor for 3D reconstruction techniques. The quality of the 3D model improves as the resolution decreases, allowing for more precise and accurate reconstructions. The resolution is primarily determined by the algorithm used for the reconstruction and the input images. Previously, the resolution of a 3D model was determined by the object's diameter and the camera's spatial acquisition rate. However, advancements in computer vision and technology have made this relationship unreliable, producing only qualitative results. Despite this, the notion that a more refined observation angle leads to higher-quality images and lower final resolution persists. Therefore, an accurate measurement of a reconstructed 3D model's resolution may indicate the quality of the reconstruction itself (van Heel et *al*., 2020).

The algorithms to measure the resolution of 3D models can be grouped in two major categories (Penczek 2002):

- techniques based on the comparison of averaged subsets of the data, such as the Fourier Shell Correlation (FSC) (van Heel *et al.* 2005) or the Differential Phase Residual (DPR) (Frank *et al.* 1981).
- algorithms based on the Fourier transform of individual images, such as the Q-factor (Kessel *et al.* 1985) and the spectral signal-to-noise ratio (SSNR) (Unser *et al.* 1987).

The first group of algorithms has a significant advantage over the second, as it can measure the resolution in both 2D and 3D (Penczek, 2002). The FSC is the dominant method used to measure resolution and has become the standard in recent years. FSC was proposed in 1987 by Harauz and Van Heel and is the 3D extension of the Fourier Ring Correlation (FRC). FSC measures the cross-correlation between two 3D models in the Fourier space created from two subsets of the same dataset. This method compares equivalent regions of the two models based on frequency and determines the resolution as the frequency at which the FSC drops below a specific threshold. The threshold is conventionally kept at 0.143, derived from the correlation between a reconstructed density map and a perfect reference map.

However, the efficacy of FSC has been widely debated due to the structural limitations introduced by the FSC itself. The ratio behind splitting the dataset to create two different models inevitably biases the final resolution, and FSC produces only a global value that does not consider all the peculiarities of the reconstructed model. To overcome these problems, the ResMap algorithm was proposed (Kucukelbir *et al.*, 2013). ResMap detects the features of a model by fitting a 3D sinusoidal function in different points of the volume and saves the wavelength of the smallest sinusoid detectable above noise.

2.3 A novel semi-automatised methodology

The 3D reconstruction methodology proposed in this paper is based on the Structure-from-Motion (SfM) technique, which reconstructs the 3D volume of an object from a series of 2D images captured from different views by a camera. Typically, SfM-based reconstruction methodologies identify feature points (also called keypoints) in the acquired images (Wu *et al.*, 2010) and match them across the entire series (Iglhaut *et al*., 2019). The matched keypoints are further refined to remove any outliers, a process commonly known as bundle adjustment. Thus, the position of the keypoints in different images is utilised to simultaneously compute the camera's pose and assign a set of 3D coordinates to each keypoint, resulting in a 3D sparse point cloud of the scene.

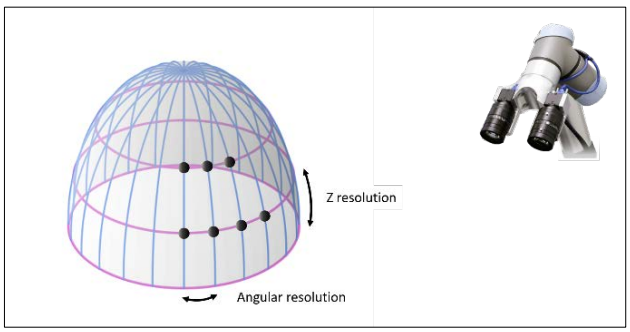

Figure 1. Schematic representation of the 2D images acquisition system. A robotic arm UR3, equipped with a stereo camera, moves along circular trajectories of variable radius, drawing a hemisphere around the object to be reconstructed. The arm stops at a pre-defined pace, enabling the camera to acquire images. The acquisition points are labeled with black dots, corresponding to intersections between the circular trajectories (in violet), representing the Z resolution, and the acquisition rate (in cornflower blue), which sets the angular resolution.

However, the quality of the reconstruction and the resolution of the reconstructed scene in SfM-based techniques are significantly affected by the number of views and the accuracy of the pose estimation. To overcome these limitations, an automated routine was designed to acquire a high number of images at a constant pace. A robotic arm UR3 equipped with an RGB camera mounted on its wrist was programmed to perform circular trajectories centred around the scene to reconstruct in 3D at different height values. The radius of these circumferences decreases with the height, creating a hemisphere around the object to be reconstructed. The number of circular trajectories and the acquisition rate define the Z resolution and the angular resolution, respectively. The robotic arm travels along circular trajectories, stopping to acquire images of the scene at a pace determined by the acquisition rate (Figure 1). The angular and Z resolution have a significant impact on the quality of the reconstructed 3D model. A higher acquisition rate results in a more resolved reconstruction. However, processing a large number of images can be computationally demanding, resulting in longer processing times and requiring more powerful workstations. Therefore, both resolution values need to be carefully selected, considering the details of the object to reconstruct, the desired final resolution, and the processing time.

To determine the best values, an AI-based algorithm has been developed. The robot rotates around the center of the scene at a fixed distance, acquiring four images 90 degrees apart at a 45-degree angle from the horizontal plane. These images are combined to obtain a coarse 3D model to estimate the object's dimensions and center of mass, which are then used as input for the AI-based technique. This technique outputs the radius of the circumferences, and the angular and Z resolutions.

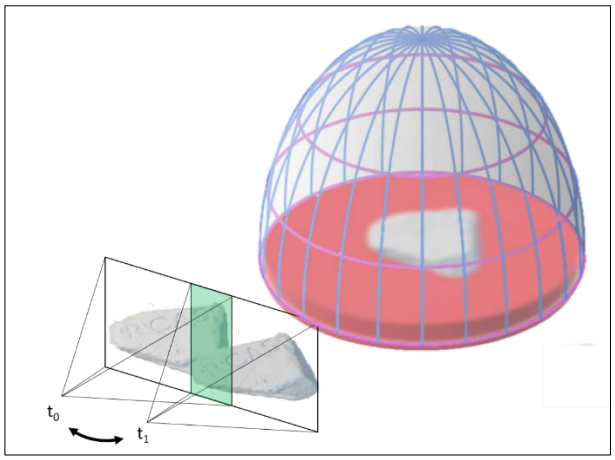

Figure 2. Schematic representation of the AI-based technique defining the camera acquisition rates. The system generates an initial coarse 3D model to estimate its dimensions and center of mass of the object, and then processes these values to determine the radius of the circumferences, as well as the angular and Z resolutions. The technique selects the best solutions by ensuring that the object does not overlap by more than 90% in consecutive images (highlighted in green).

The system identifies the best solutions by ensuring that there is no more than 90% overlapping of the object in consecutive images (Figure 2).

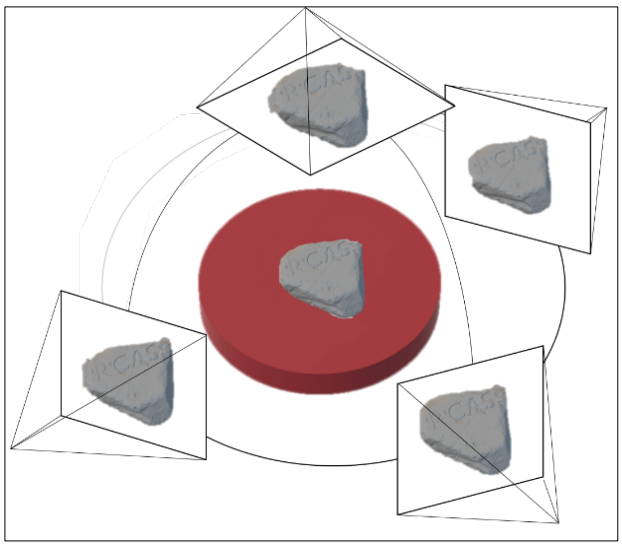

Figure 3. Schematic representation of the reconstruction process. The robotic arm equipped with the vision system rotates around the center of the scene (the red pedestal) following pre-defined circular trajectories and acquires a series of images.

The ability to finely adjust both the Z and angular resolution values allows for a precise and dense acquisition of views, resulting in high-accuracy reconstruction of the scene. Additionally, these values can be easily modified to optimise the reconstruction for objects of different sizes. The pose of each view is determined by the pose of the tool center point, which is expressed as two sets of 3 coordinates to model its position and orientation, respectively (refer to Figure 3).

Another advantage provided by the robotic arm is its ability to move with millimetre precision, thus overcoming one of the main limitations of the SfM methodology. By providing highly accurate pose values and eliminating operator errors in the workflow, the poses of all views are well constrained, simplifying and improving the performance of the reconstruction algorithm. The poses obtained are first used to correct any camera distortion, and then to assign a set of 3D coordinates to the different images, generating a dense and refined point cloud output.

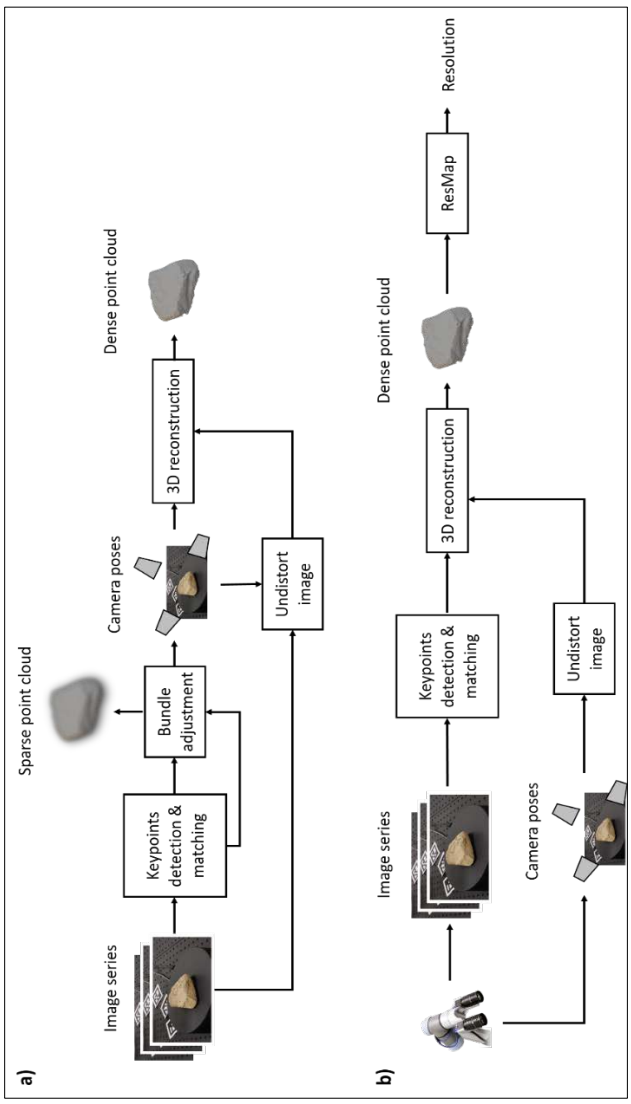

Figure 4. Schematic representation of the comparison between the conventional 3D reconstruction technique (**a**) and the one proposed in this paper (**b**). The proposed method uses a robot to rotate around the scene, which constrains and simplifies the methodology.

To evaluate the quality of the reconstruction, it is essential to measure the resolution of the 3D model. Therefore, the resulting 3D model obtained by the proposed methodology is processed using ResMap algorithm. This algorithm produces a local resolution map and associates a distribution of values to the resolution of the density map. A precise local estimation of the resolution enables structural analysis, setting a limit to the significant elements in the 3D reconstruction. Furthermore, the robustness of this methodology to noise helps avoid confusing high-quality results with high-frequency noise, as the latter may be visually appealing but can deteriorate the information (Röhrbein *et al.*, 2015).
Figure 4 illustrates a comparison between the traditional methodology(a) and the one proposed in this paper (b), which eliminates the need for bundle adjustment and generates directly a dense point cloud. This approach not only improves efficiency but also reduces computational overhead, resulting in faster reconstruction times during the computation stage.

3. CONCLUSIONS

Acquiring images and transforming them into 3D models is a complex process that requires careful consideration of various factors. One of the most critical factors is the acquisition rate of the images, which can significantly impact the accuracy and consistency of the final reconstruction. Conventional methodologies often involve capturing images in an unordered sequence, which can lead to variations in the reconstructed 3D models, even when using the same image sequence.

To address this issue, our proposed method utilises a robotic arm to standardise the image acquisition process. The motion of the robotic arm is programmed to move in an optimised manner, capturing images at specific intervals by avoiding redundant information, high computational times, and ensuring high-quality results. This approach offers several advantages over traditional methodologies. For one, it increases the repeatability and consistency of the reconstructions, as well as improving their accuracy and reliability.

A key advantage of using a robotic arm for image acquisition is that it enables us to constrain the acquisition rate. By capturing images at a consistent pace, we can obtain a more accurate representation of the scene, even when there are small variations due to degradation or damage. Because the images are captured at a consistent interval, the impact of any changes in the scene is minimised, resulting in a more accurate representation of the scene.

Another advantage of our proposed method is that it provides a more reliable basis for comparing different 3D models. With highly consistent acquisition rates, we can be confident that any differences between the models are due to actual changes in the scene, rather than variations in the image sequence or reconstruction process. This increases the accuracy of the final model and makes it easier to identify any changes or differences between the models.

Our proposed method offers promising opportunities for the future development of 4D applications in cultural heritage preservation. By adding time as the fourth dimension, we can track changes over time, which is crucial for the conservation and preservation of material culture, and enables the creation of more advanced Digital Twins beyond simple 3D scans. Monitoring structural variations affecting the reconstructed volumes of the scene through time enhances the precision of monitoring these transformations. This has become an essential component of conservational maintenance for archaeological artifacts, enabling better-informed decisions about how to protect and preserve these valuable pieces of our collective cultural heritage.

Using our method, mistakes made during the evaluation of modifications are minimised, providing more reliable distinctions in conditions of damage or aging. This is particularly important in the field of cultural heritage preservation, where accurate and reliable data is crucial for making informed decisions about how to conserve and protect these valuable artifacts.

REFERENCES

Chandler, J. H., Buckley, S., 2016. Structure from motion (SFM) photogrammetry vs terrestrial laser scanning. In: Carpenter, M. B., Keane, C. M. (eds), *Geoscince Handbook 2016: AGI Data Sheets*, 5th ed. Alexandria, VA: American Geosciences Institute, Section 20.1.

Ćosović, M., Maksimović, M., 2022. Application of the digital twin concept in cultural heritage. In: Amelio, A., Montelpare, S., Ursino, D. (eds.) *Proceedings of the 1st International Virtual Conference on Visual Pattern Extraction and Recognition for Cultural Heritage Understanding, 12 September 2022.*

Europeana Foundation 2023. Common European data space for cultural heritage, accessed 28 April 2023, https://pro.europeana.eu/page/common-european-data-space-for-cultural-heritage

Frank, J., Verschoor, A., Boublik, M., 1981. Computer averaging of electron micrographs of 40s ribosomal subunits. *Science*, 214(4527), 1353–1355.

Gabellone, F. 2022. Digital Twin: a new perspective for cultural heritage management and fruition. *Acta IMEKO* 11(1).

Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., Rosette, J., 2019. Structure from Motion Photogrammetry in Forestry: a Review. *Current Forestry Reports* 5, 155-168: doi.org/10.1007/s40725-019-00094-3

Kaneda, A., Nakagawa, T., Tamura, K., Noshita, K., Nakao, H., 2022. A proposal of a new automated method for SfM/MVS 3D reconstruction through comparison of 3D data by SfM/MVS and handheld laser scanners. *PLoS One*, 17(7): e0270660. Doi.org/10.1371/journal.pone.0270660

Kessel, M., Radermacher, M., Frank, J., 1985. The structure of the stalk surface layer of a brine pond microorganism: correlation averaging applied to a double layered lattice structure. *Journal of Microscopy*, 139(1):63–74.

Kock, R., May, S., Nuchter, A., 2017. Detection and purging of specular reflective and transparent object influences in 3D range measurements. *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, XLII/2/W3, 377-384.* doi.org/10.5194/isprs-archives-XLII-2-W3-377-2017

Kucukelbir, A., Sigworth, F. J., Tagare, H. D. 2013. Quantifying the local resolution of cryo-EM density maps. *Nature Methods*, 11(1):63–65.

Luther, W., Baloian, N., Biella, D., Sacher, D., 2023. Digital Twins and Enabling Technologies in Museums and Cultural Heritage: An Overview. *Sensors* 23(3): 1583. Doi.org/10.3390/s23031583

Penczek, P. A., 2002. Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms. *Journal of Structural Biology*, 138(1-2):34–46.

Polo, M., Felicísimo, Á. M., Durán-Domínguez, G., 2022. Accurate 3D models in both geometry and texture: An archaeological application. *Digital Applications in Archaeology and Cultural Heritage*, 27. doi.org/10.1016/j.daach.2022.e00248

Rachakonda, P. K., Muralikrishnan, B., Sawyer, D., 2019. Sources of errors in structured light 3D scanners. *Proceedings SPIE 10991, Dimensional Optical Metrology and Inspection for Practical Applications VIII.* doi.org/10.1117/12.2518126

Remondino, F., 2011. Heritage Recording and 3D Modeling with Photogrammetry and 3D Scanning. *Remote Sens.*, 3, 1104-1138. Doi.org/10.3390/rs3061104

Röhrbein, F., Goddard, P., Schneider, M., James, G., Guo, K., 2015. How does image noise affect actual and predicted human gaze allocation in assessing image quality? *Vision research* 112: 11-25.

Tomasi, C., Kanade, T., 1992. Shape and motion from image streams under orthography: a factorization method. *International journal of computer vision*, 9:2, 137-154. doi.org/10.1073/pnas.90.21.9795

Tsvetaeva, A., 2022. Museums and Digital Technologies: To what extent can digitization of museums collections help access, promote and preserve cultural heritage? Case studies on Mauritshuis and Kunstmuseum, The Hague.

Unser, M., Trus, B. L., Steven, A. C., 1987. A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy*, 23(1):39–51.

Van Heel, M., Gowen, B., Matadeen, R. M. Orlova, E. V., Finn, R., Pape, T., Cohen D., Stark, H., Schmidt, R., Schatz, M., Patwardhan, A., 2000. Single-particle electron cryo-microscopy: towards atomic resolution. *Quarterly Reviews of Biophysics*, 33(4):307–369.

Van Heel, M., Schatz, M., 2005. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3):250–262.

Wu, X., Shi, Z., Zhong, Y. 2010. Detailed analysis and evaluation of keypoint extraction methods. *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. Vol. 2. IEEE.

Yoo, J. C., Han, T. H., 2009. Fast normalized cross-correlation. *Circuits, Systems and Signal Processing*, 28(6):819–843.