GA – PROJECT NUMBER:	**101158046**
PROJECT ACRONYM:	AUTOMATA
PROJECT TITLE:	AUTOmated enriched digitisation of Archaeological liThics and cerAmics
CALL/TOPIC:	HORIZON-CL2-2023-HERITAGE-ECCCH-01-02
TYPE OF ACTION	HORIZON RIA
PRINCIPAL INVESTIGATOR	Prof Gabriele Gattiglia, UNIPI
TEL:	+39 050 2215228
E-MAIL:	gabriele.gattiglia@unipi.it

This project has received funding from the European Union's HORIZON RIA research and innovation programme under grant agreement N. 101158046

D 5.3 Reference enriched 3D data

Version: 1.0

Revision: first release

Work Package: 5 - Technologies for enriched digitisation
Lead Author (Org): Arthur Leck (UBM), Clément Joubert (INRIA)
Contributing Author(s) (Org): Rémy Chapoulie (UBM), Romain Pacanowski (INRIA), Nevio Dubbini (MIN), Daniël Van Helden (KCL), Bruno Dutailly (UBM), Maxime Constant (UBM), Gabriele Gattiglia (UNIPI), Martina Naso (UNIPI)
Due Date: M22
Date: 21/07/2026

Dissemination Level		
P	**Public**	**X**
C	**Confidential, only for members of the consortium and the Commission Services**	

Revision History

Revision	Date	Author	Description
0.1	02/07/2026	Clément Joubert (INRIA)	Structure and draft content
0.2	08/07/2026	Clément Joubert, Romain Pacanowski (INRIA)	Draft content for 3.1
0.3	10/07/2026	Arthur Leck (UBM)	Draft content
0.4	13/07/2026	Nevio Dubbini (MIN)	Draft content for section 5
0.5	13/07/2026	Clément Joubert (INRIA)	Draft content for section 2
0.6	14/07/2026	Daan van Helden (KCL)	Edit section 5, minor edits, renumber figures
0.7	15/07/2026	Arthur Leck (UBM)	Draft content, section 4
0.8	16/07/2026	Arthur Leck (UBM)	Draft content, section 3
0.9	16/07/2026	Gabriele Gattiglia	Draft revision
1.0	16/07/2026	Clément Joubert (INRIA)	Draft content, conclusion
1.1	17/07/2026	Arthur Leck (UBM)	Draft revision
1.2	17/07/2026	Clément Joubert (INRIA)	General revision
1.3	20/07/2026	Nevio Dubbini (MIN)	Draft content, general revision
1.4	20/07/2026	Arthur Leck, Bruno Dutailly, Maxime Constant (UBM)	RIS3D server and application build, Draft content for section 4
1.5	21/07/2026	Gabriele Gattiglia, Martina Naso (UNIPI) Nevio Dubbini (MIN), Arthur Leck, Remy Chapoulie(UBM), Clément Joubert (INRIA)	Final revision

Disclaimer

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Abbreviations

WP: Work Package

M: Month

UNIPI: Università di Pisa

UBM: Université Bordeaux Montaigne

UoY: University of York

INRAP: Institut National de Recherches Archéologiques Préventives

INRIA: Institut national de recherche en sciences et technologies du numérique

AMZ: Arheoloski Muzej u Zagrebu

QB: QBrobotics Srl

HUJ: The Hebrew University of Jerusalem

MIN: Miningful srls

KCL: King's College London

IIT: Fondazione Istituto Italiano di Tecnologia

UB: Universitat de Barcelona

CL: Culture Lab

Executive summary

This deliverable presents the main results of Work Package 5 and the implementation of a software-supported workflow for generating enriched 3D models of archaeological artefacts. The principal result of D5.3 consists of the software components and reference datasets made available with this document, while the report provides the accompanying methodological and technical documentation. Building upon the ontology and metadata framework (D5.1) and the dedicated 3D database (D5.2), the work delivered in D5.3 includes the creation of reference enriched datasets and the methods used to register archaeometric measurements with high-resolution 3D models.

A key software component of this workflow is the adaptation of the **Reference Information System in 3D (RIS3D)** to the requirements of the AUTOMATA project. RIS3D provides an interactive platform for visualising, managing and annotating enriched 3D models, enabling hyperspectral images, pXRF and Raman measurements to be spatially linked to the 3D surface. The materials and code created for the implementation of this deliverable are available in openly accessible repositories:
(i)the RIS3D application, together with the initial collection of enriched 3D models and associated multimodal datasets, can be downloaded and accessed through the AUTOMATA RIS3D platform at https://automata.3dris.science/. Access and installation instructions are provided in Section 4.2
(ii) the code developed for the automatic selection and visualisation of analytical target locations is available at https://github.com/minstudio/automata_crosshair.
The source code of RIS3D and the AUTOMATA-specific software adaptations will be available with Deliverable 6.2 *Visualisation and Interaction* on M28.

The document also includes the automated workflows developed to integrate multimodal archaeometric data into enriched 3D models. These include an automated photogrammetric reconstruction pipeline, hyperspectral image registration, and the automatic detection of Regions of Interest (ROIs) to support the selection of target locations for pXRF and Raman analyses. Together with the calibration monitoring, quality assessment, and adaptive acquisition framework presented in D4.1, these developments establish the methodological foundations for the AUTOMATA closed-loop digitisation workflow.

Section 1 introduces the document and outlines the objectives of this deliverable. Section 2 describes the photogrammetric acquisition setup and parameters. Section 3 presents the photogrammetric and hyperspectral registration workflow, which automatically aligns hyperspectral images with the 3D model. This is a key step in the creation of enriched 3D models. Section 4 introduces the Reference Information System in 3D (RIS3D), explains how it operates, and provides guidance on how to use it to access and explore enriched 3D models. Section 5 describes the automatic detection of Regions of Interest (ROIs) for XRF and Raman analyses. Finally, Section 6 concludes the document and outlines future developments related to the registration and visualisation of enriched 3D models.

1 Introduction

Work Package 5 (WP5) focuses on the development of the methodological and software foundations of a workflow for the generation of enriched 3D models of archaeological artefacts. Building upon high-resolution 3D digitisation, the objective is to integrate complementary archaeometric information, including elemental, mineralogical and spectral analyses, into a common spatial framework. Beyond the production of digital replicas, the work aims to create reference datasets in which geometry, appearance and analytical measurements are spatially linked and fully documented. These enriched datasets provide the foundation for the development, training and validation of the automated workflows developed throughout the AUTOMATA project.

The first two deliverables of WP5 established the framework required for this objective. Deliverable D5.1 defined the ontology and metadata scheme necessary to ensure interoperability, traceability and FAIR management of enriched 3D data. Deliverable D5.2 presented the implementation of the dedicated 3D database, providing the infrastructure for storing, managing and sharing enriched models together with their associated metadata. Together, these deliverables established the data model and digital infrastructure upon which the present work is built.

Deliverable D5.3 focuses on the generation of reference enriched 3D datasets and concludes the activities of WP5. It addresses the practical implementation of the enrichment workflow by describing the acquisition protocols, data processing methods and software tools developed to associate archaeometric measurements with 3D models. The work covers both the manual generation of reference datasets, required for the creation of reliable ground-truth data, and the implemented procedures supporting the automation of the enrichment process.

Particular attention is given to the spatial registration of multimodal data acquired using complementary analytical techniques, including portable X-ray fluorescence (pXRF), Raman spectroscopy and hyperspectral imaging. Since these measurements are collected independently from the 3D acquisition process, reliable spatial registration is essential to associate each analytical observation with its corresponding location on the digital model. To address this challenge, dedicated acquisition supports, photogrammetric registration procedures and data integration tools have been developed. The resulting workflows enable analytical measurements to be projected onto the 3D surface while preserving their spatial relationships and associated metadata.

This deliverable also presents the AUTOMATA implementation of the Reference Information System in 3D (RIS3D), a software platform designed to visualise, manage and annotate enriched 3D models. RIS3D provides an interface for displaying hyperspectral imagery and recording archaeometric data on the model's surface. This application is distributed together with an initial set of enriched 3D models and associated HSI, pXRF and Raman spectrometry datasets.

D5.3 also includes the automated workflows developed for the generation of enriched 3D models and the integration of multimodal archaeometric data. Building upon the manually created reference datasets required for ground-truth generation, the proposed pipeline automates photogrammetric reconstruction, hyperspectral image registration, and the spatial enrichment of 3D models, thereby reducing the number of manual processing steps while maintaining the geometric accuracy and traceability required for archaeological documentation. In conjunction with the calibration monitoring, quality assessment, and adaptive acquisition framework presented in Deliverable D4.1, these developments establish the methodological basis for an increasingly automated acquisition pipeline, in which validated multimodal data are progressively integrated into enriched 3D reference models. As an integral component of this automated workflow, this deliverable includes the automatic detection of Regions of Interest (ROIs) on archaeological artefacts. By combining geometric information extracted from the 3D models with spectral information derived from hyperspectral imaging, the proposed approach automatically identifies the most relevant areas for archaeometric investigation. These ROIs are subsequently projected onto the hyperspectral images and registered on the corresponding 3D models, defining candidate target locations for targeted pXRF and Raman analyses. Together with the diagnostic and feedback mechanisms developed in D4.1, this workflow contributes to a closed-loop digitisation strategy in which acquisition, validation, and spatial integration are progressively coupled within a unified automated framework.

The selected target locations and their spatial coordinates are structured so that they can be transferred to the robotic control workflow and associated with the corresponding measurements on the 3D model. This provides the basis for traceable integration of archaeometric analyses within the future fully automated system.

These developments constitute an important step towards the fully automated acquisition and enrichment pipeline targeted by the AUTOMATA project.

2 Acquisition setup

This section describes the acquisition setup developed within the project to enable the integration of geometric, hyperspectral and archaeometric data. It presents the preparation of the existing 3D models, the design of the acquisition support, and the constraints considered to ensure reliable photogrammetric reconstruction and hyperspectral registration.

2.1 3D data processing

The objects investigated in this work originate from previous digitisation campaigns carried out within the project (see D5.2, section 2). Existing 3D models were therefore used as the geometric reference for the integration of the newly acquired archaeometric data. Before their integration into the RIS3D platform, these models underwent a preparation step to improve their geometric quality and compatibility with subsequent processing. The preparation consisted primarily of manually cleaning the meshes by removing reconstruction artefacts, isolated components, and topological defects (Figure 1). Particular attention was paid to producing watertight and manifold meshes, ensuring that the models could be reliably visualised, processed, and enriched with additional information in the following stages of the workflow.

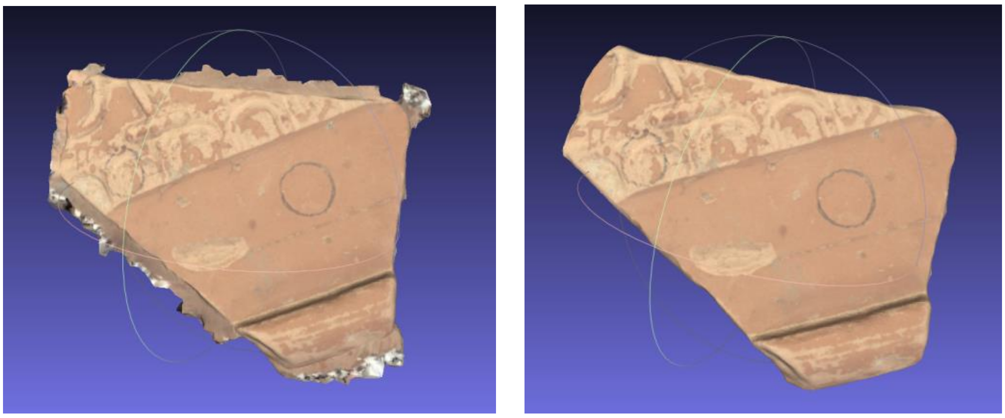

Figure 1 - 3D model of artefact Lot2n11 generated with Agisoft Metashape: original reconstruction (left) and manually cleaned mesh (right).

Although 3D models from previous acquisition campaigns were available, the associated photogrammetric and hyperspectral datasets had not been acquired under the same experimental conditions. In particular, the object was repositioned between the two acquisitions (Figure 2), making the registration of the hyperspectral data with the photogrammetric reconstruction more challenging and less reliable. These earlier datasets were therefore retained as test material for RIS3D but were not suitable for validating the automatic registration workflow. To generate a consistent reference dataset for D5.3, a new acquisition setup was developed in which both the photogrammetric and hyperspectral images are captured sequentially without moving the object. This protocol ensures a consistent acquisition geometry and supports more reliable registration of the different data modalities described in the following sections.

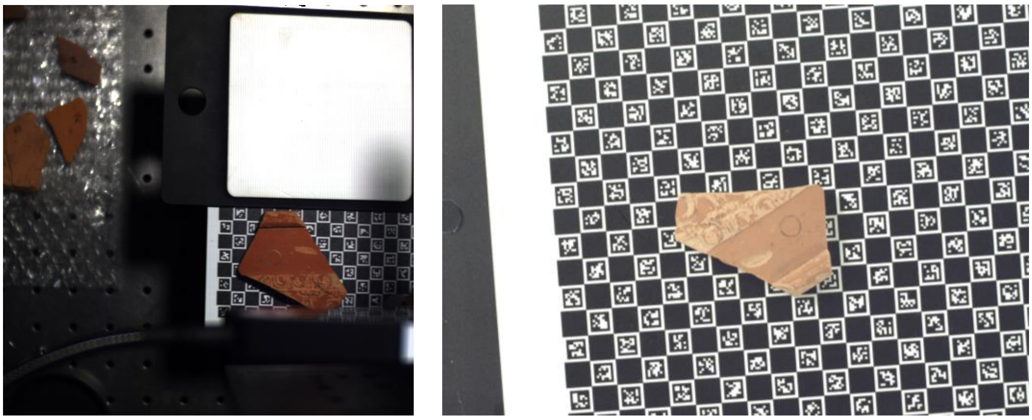

Figure 2 - Hyperspectral image (left) and photogrammetric image (right) of artefact Lot2n11, acquired in different conditions during the earlier campaign.

2.2 Support design for photogrammetry and registration

Following numerous test configurations, some of which are described in Deliverable D5.2, we established a manually operated acquisition protocol (see Figure 3) that reproduces the future fully automated acquisition process closely and provides a test configuration for subsequent transfer to the robotic-arm system developed within the AUTOMATA project. This dedicated acquisition setup was developed to ensure the consistent registration of photogrammetric and hyperspectral images. Since both modalities must ultimately share the same geometric reference system, the RGB images used for photogrammetric reconstruction are acquired under the same conditions and from viewpoints compatible with those used for hyperspectral acquisition.

Figure 3 - HSI and photogrammetry set-up. HinaLea VNIR 4250 hyperspectral camera (left), Specim IQ hyperspectral camera (centre) and Nikon Z7 II camera for photogrammetry (right).

Image acquisition was carried out using a Nikon Z7 II camera equipped with a Nikon 24–70 mm f/2.8 lens. Illumination was provided by a Godox AR400 ring flash. Several support configurations were evaluated during the development of the acquisition protocol, including nylon strings and transparent stands. These solutions expose both the upper and lower parts of the object while reducing the amount of occluded surface visible to the cameras. However, each solution introduces its own limitations. Transparent supports may generate reflections that negatively affect photogrammetric reconstruction, while thicker supports can occlude portions of the artefact and reduce the completeness of the reconstructed model. The current acquisition setup uses a circular support made of poly(methyl methacrylate) (PMMA), a transparent thermoplastic often referred to by the trade name Plexiglas (Figure 4). The structure holding the support was mounted on a rotating table to remain stable then taking pictures on the azimuthal angle. This support was selected as the most suitable configuration among those tested in terms of object stability, surface visibility, and ease of reconstruction.

This PMMA sample holder is disc-shaped, 4 mm thick, and 30 cm in diameter, mounted on a motorised rotation stage. This configuration was chosen to reproduce, as closely as possible, the acquisition conditions of the future automated system.

Figure 4 - Circular PMMA support used for the reference acquisitions: bottom view with visual patterns for photogrammetry (left), and upper surface of the Plexiglas support.

Forty circular coded photogrammetric targets from Agisoft Metashape, with a diameter of 1 cm each were attached directly to the transparent support so that they remained visible from both sides. The targets were printed on both sides of a single sheet of paper, with their centres precisely aligned on both faces, allowing the photogrammetric software to identify the same reference points from the top and bottom views. In addition, eight small 3D-printed mounts were fixed to each side of the disc to support extra coded targets, ensuring that a sufficient number of targets remained visible even at the most constrained viewing angles. Two sizes of mounts were used. The first type was designed to hold 1 cm diameter targets, while the second, smaller type was designed to hold 0.6 cm diameter targets in order to minimise the obstruction caused by the mounts. The successful detection of the 0.6 cm diameter targets suggests that the circularly arranged targets could also be reduced in size in future iterations. Feature-based photogrammetric methods estimate the camera's poses by identifying common visual patterns across multiple images. Consequently, static markers are placed around the acquisition area to facilitate feature matching and improve the robustness of the camera pose estimation, as explained in the next section.

Each acquisition consisted of approximately 216 photographs, representing a clear oversampling of the object, providing high image overlap for robust 3D reconstruction. Images were acquired from eight zenithal positions, four above the object and four below the support, combined with approximately 24 azimuthal positions distributed over a rotation of the sample. The camera viewpoints covered an angular range of approximately ±45 ° relative to the horizontal plane.

3 Photogrammetric and hyperspectral registration workflow

This section describes the workflow developed to generate a photogrammetric 3D reconstruction and to register hyperspectral information within the same spatial reference system. The objective is to establish a reliable spatial link between the geometric information provided by photogrammetry and the target location of pXRF and Raman analyses. Since the Raman and XRF measurement locations are spatially determined from hyperspectral data, registering the hyperspectral image with the reconstructed 3D model is a key requirement for aligning archeometric measurements with their corresponding position on the artefact surface.

The proposed workflow combines automated sparse reconstruction, hyperspectral image registration, view-dependent mask generation and dense 3D reconstruction. Its principal outputs are the textured 3D model and the parameters required to project the hyperspectral data onto it. The individual processing steps, from image acquisition to the generation of the final textured model, are described in the following sections.

The earlier datasets produced within the project were acquired independently for photogrammetry, HSI, pXRF and Raman analysis, with the artefact repositioned between acquisition stages. These datasets could therefore be used to test visualisation and data association in RIS3D, but they were not suitable for reliable automatic HSI-to-3D registration. To validate the registration procedure, a new, smaller reference dataset was acquired using the fixed-object protocol described in Section 2, in which the artefact remained in the same position during both photogrammetric and HSI acquisition.

3.1 Automatic photogrammetric pipeline

With the increasing interest in image-based 3D reconstruction, both the diversity of available methods and the volume of related research continue to grow rapidly. The cultural heritage community has adopted this evolution by exploring both established and emerging approaches, such as Gaussian Splatting (Fei et al., 2024) and NeRF (Yao et al., 2026), for the 3D reconstruction and visualisation of archaeological objects and sites. These methods have demonstrated remarkable capabilities for producing highly realistic renderings and interactive visualisations. However, they are primarily designed for novel-view synthesis rather than accurate metric reconstruction. In archaeological applications where geometric precision, repeatability, and computational efficiency are essential, conventional photogrammetric workflows remain the preferred solution and continue to provide reliable geometric information and a favourable time-computation trade-off (Maté-González, 2025). Consequently, we adopt the photogrammetric workflow for this study (Figure 5).

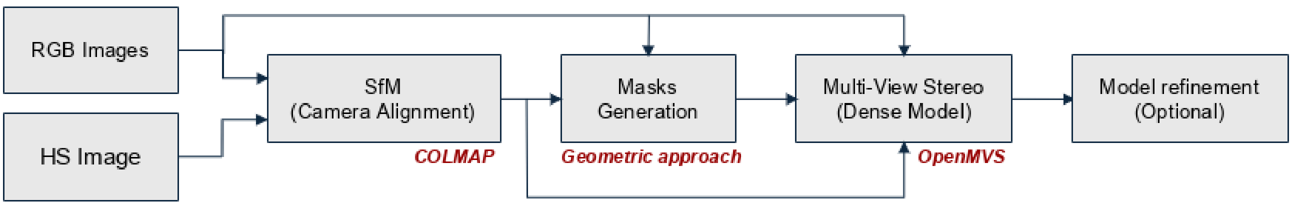

Figure 5 - Automatic photogrammetric pipeline

The primary objective of the proposed photogrammetric workflow is to accurately estimate camera poses and establish reliable correspondences among the acquired images, thereby enabling the reconstruction of a consistent 3D model. The same registration principle is subsequently applied to align the hyperspectral images with the photogrammetric dataset. For that reason, it is important to define the photogrammetric pipeline and each step involved in computing the final 3D model. To ensure transparency and reproducibility, the entire workflow relies exclusively on two open-source software packages: COLMAP (Schönberger et al., 2016) for image registration and sparse reconstruction and OpenMVS (Cernea, 2020) for dense reconstruction and textured mesh generation. In addition to being open source, an essential requirement for a European-funded project, COLMAP supports incremental photogrammetry. This makes it possible to perform photogrammetric reconstruction during image acquisition, thereby reducing the overall duration of the AUTOMATA workflow.

3.2 Image acquisition and hyperspectral image registration

As described in D5.2, the photogrammetric images are acquired following a spherical trajectory around the object. This acquisition strategy provides high image overlap and a wide range of viewing angles, which are required for robust image matching and accurate 3D reconstruction. The acquired RGB images are first converted to JPEG format and then imported into COLMAP for feature extraction, image matching, camera pose estimation, and sparse reconstruction (Figure 4). Since COLMAP operates on RGB images, the hyperspectral data must first be converted into a compatible representation. For this purpose, a pseudo-RGB image is generated from the hyperspectral data by creating a linear reconstruction with three representative spectral bands corresponding to the visible red, green, and blue channels. For the SPECIM IQ hyperspectral camera used in this deliverable, the selected bands correspond to the default RGB visualisation wavelengths provided by the system: 598.60 nm for the red channel, 548.55 nm for the green channel and 449.35 nm for the blue channel. The resulting pseudo-RGB image is then processed alongside the photogrammetric images in COLMAP to register the hyperspectral data within the same photogrammetric reference frame (Figure 6). This registration step requires sufficient visual overlap between the hyperspectral pseudo-RGB image and RGB photogrammetric images. Therefore, the acquisition protocol must ensure that at least one RGB image is captured from a viewpoint close to the hyperspectral camera position. This shared viewpoint maximises the common features between the two cameras, increasing the number of corresponding features detected by the SIFT algorithm and, consequently, improving the robustness of the feature matching process.

*Figure 6 - **Left**: Pseudo-RGB image generated from SPECIM IQ hyperspectral data (i.e. a 'regular' image extracted from the hyperspectral data). **Right**: RGB image from Nikon Z7 II camera for the object Lot1_112.*

3.3 Incremental Structure-from-Motion reconstruction

COLMAP relies on the Scale-Invariant Feature Transform (SIFT) algorithm (Lowe, 1999) to detect and describe distinctive features in the images. These features are then matched between overlapping images to identify common points observed from different viewpoints (Figure 7). The resulting correspondences are used to estimate the relative position and orientation of the cameras and to progressively reconstruct the 3D structure of the scene. The quality of this feature-matching step is essential for the successful registration of the hyperspectral images. If an insufficient number of reliable correspondences is found between the hyperspectral pseudo-RGB image and the photogrammetric images, the alignment between both cameras cannot be achieved. Therefore, objects with limited visual information, i.e., poor texture, can represent a challenge for automatic registration. Ensuring the presence of recognisable visual patterns and sufficient texture on the object is essential to ensure robust feature matching, accurate registration, and precise sparse reconstruction.

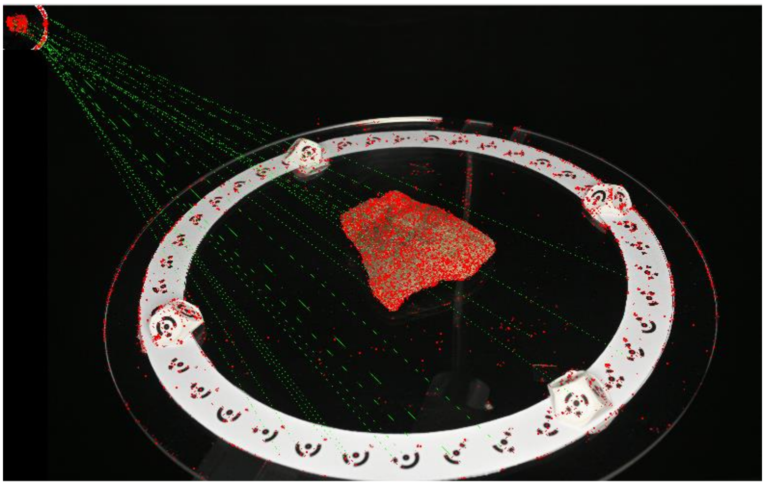

Figure 7 - Example of feature extraction and matching performed with COLMAP on Lot1_112 with the two images of Figure 6. The hyperspectral image is inset in the upper left, and the photogrammetric image is on the right. Each red dot is a SIFT feature, and each green line represents a feature match between the two images.

Following the feature extraction and matching, COLMAP applies an incremental Structure-from-Motion (SfM) algorithm to estimate the camera poses and generate a sparse point cloud (Figure 8). The incremental approach starts from an initial set of registered images and progressively adds new images while refining the 3D reconstruction. At each iteration, new observations are triangulated, bundle adjustment is performed, and inconsistent measurements are removed. Bundle adjustment is a key optimisation step that jointly refines the estimated camera parameters, camera poses, and reconstructed 3D points. Although incremental SfM generally requires more processing time than global SfM approaches, the iterative refinement and outlier filtering provide a more robust estimation of the scene geometry, which is particularly important for our archaeological applications.

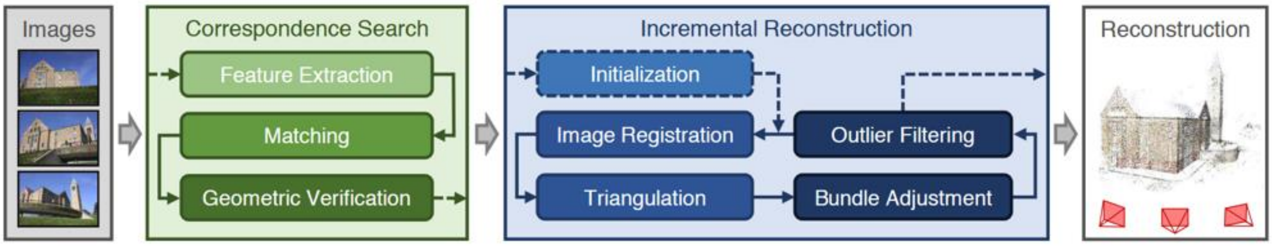

Figure 8 - Incremental Shape-from-Motion pipeline of COLMAP taken from Schönberger et al., 2016

Applying this pipeline to the fixed-object acquisition (Figure 6) produces a sparse point cloud combined with the camera positions (Figure 9). With the positions and optical parameters computed by COLMAP for the hyperspectral camera, each spectral channel of the hyperspectral cube can be projected onto the final 3D model, enabling visualisation of hyperspectral information within the photogrammetric reference frame. However, the sparse reconstruction generated by COLMAP does not provide sufficient geometric resolution for the final 3D model. Therefore, an additional dense reconstruction step is required to densify the point cloud before mesh generation and texturing. Nevertheless, the resulting point cloud from the sparse reconstruction gives useful information about the object position in photogrammetric space.

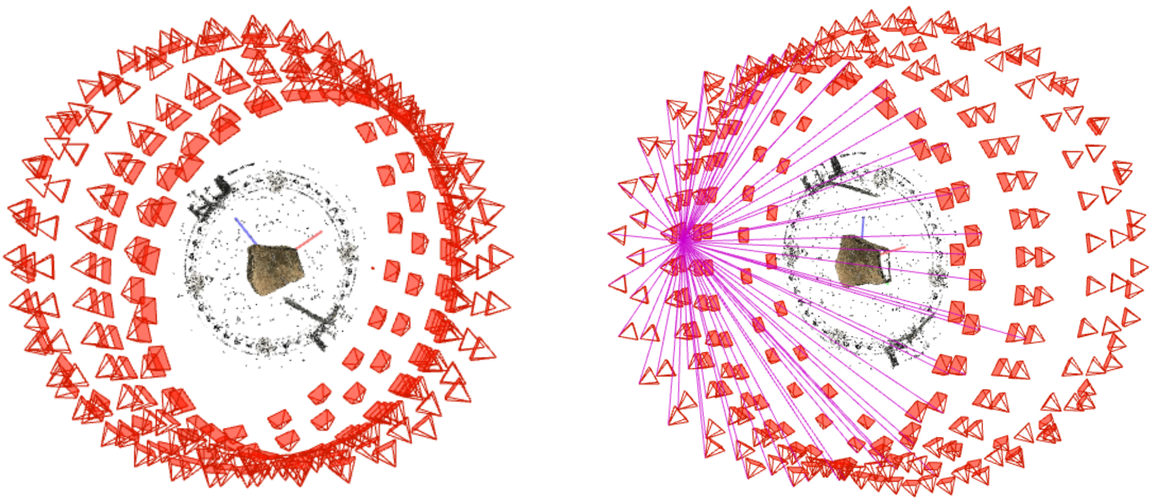

Figure 9 - Two views of COLMAP visualisation of Lot1_112 with the point cloud in the centre, the camera positions with the red rectangles, and the purple lines (right) show the hyperspectral image registration with RGB images.

3.4 Mask generation

Before performing dense reconstruction, an automatic masking procedure is applied to restrict the processing area to the object of interest. The main idea is to apply 2D masks, defined in image space, which determine the object boundaries before densifying the point cloud (cf. Figure 5). This step aims to improve the quality of the final reconstruction by preventing the generation of unnecessary points and triangles from the background and to provide a cleaner input for the following reconstruction stages. It also reduces computational cost by limiting image processing to the pixels corresponding to the object surface. The robotic acquisition system provides an estimation of the object position (D3.2) as well as its bounding box. A registration step is required to find the transformation between the photogrammetric and the robot coordinate frames, such as using the Umeyama algorithm (Umeyama, 1991) with camera positions in both coordinate systems. In the future integrated robotic workflow, the object position and bounding box will be supplied by the robotic perception system described in D3.2, following transformation between the robot and photogrammetric coordinate frames. In the D5.3 implementation, however, the initial 3D bounding box is defined manually. This bounding box is used to remove points located outside the object from the sparse point cloud generated by COLMAP (Figure 10).

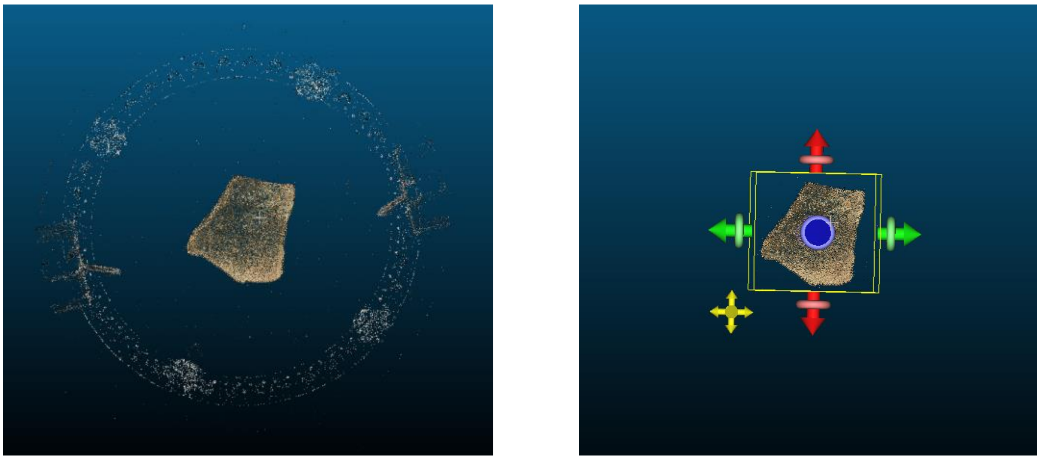

Figure 10 - Simulating the bounding box cropping with CloudCompare before (left) and after (right) cropping.

The filtered sparse point cloud is then converted into a preliminary mesh using OpenMVS. Although this mesh does not represent the final high-resolution reconstruction, it provides an approximate geometric representation of the object used to generate view-dependent masks. For each camera viewpoint, the camera poses and intrinsic parameters estimated by COLMAP are used to render the object from the corresponding position. Depth maps are generated using pyrender (Matl, 2019), a lightweight physically-based rendering framework. On these maps, every pixel that is outside of the object is equal to 0, and it is superior to 0 on the object. Then, to convert a depth map to a binary mask of the object, we apply a threshold of 0. These 2D masks are subsequently used during the dense reconstruction step to ensure that only the relevant regions of the images are processed (Figure 11). Additionally, to compensate for potential uncertainties in the sparse reconstruction or camera estimation, a morphological dilation can be applied to the generated 2D masks. This allows a safety margin around the object boundaries while preserving the advantages of the automated segmentation approach.

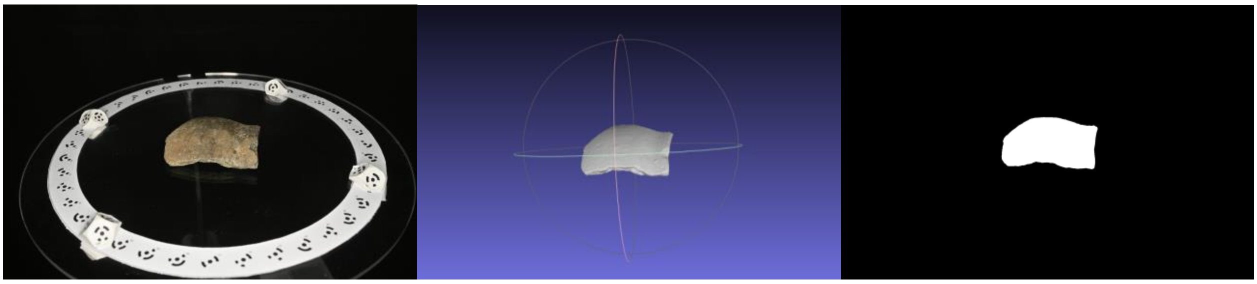

*Figure 11 - Creating a mask for each camera viewpoint from the image (**left**) with the sparse mesh of the object (**centre**) to finally obtain the mask for this viewpoint (**right**)*

3.5 Multi-view stereo

The final stage of the photogrammetric workflow consists of densifying the sparse reconstruction obtained from the Structure-from-Motion step and generating the final 3D model. This process is performed using Multi-View Stereo (MVS), which aims to recover a dense representation of the object surface from multiple images acquired from different viewpoints. In this project, OpenMVS (Cernea, 2020) is used to perform the dense reconstruction. The software implements a geometric approach based on the PatchMatch Stereo algorithm (Bleyer, 2011). The method estimates depth maps for each camera viewpoint by exploiting the correspondence between pixels observed in multiple images. These depth maps are then fused to generate a dense point cloud representing the object surface. The camera parameters and poses estimated by COLMAP are used as input for OpenMVS, together with the RGB images and the previously generated masks. These masks ensure that the dense reconstruction process focuses only on the object of interest and prevent the generation of background reconstruction defects.

Once the dense point cloud has been generated, it is converted into a surface mesh using Delaunay triangulation implemented in OpenMVS. The resulting mesh is then textured using the original RGB images to obtain a realistic representation of the reconstructed object (Figure 12). The final textured model provides the geometric reference required for the integration of hyperspectral information and the subsequent visualisation of archaeometric data within the 3D model. The outputs of this process are:

- A textured 3D model (.glb)
- Two text files that contain the intrinsic and extrinsic parameters of the hyperspectral camera in the COLMAP format (.txt)

With this pipeline, 3 different samples of ceramics were reconstructed, and we kept the original resolution of the images, which is 8256x5504.

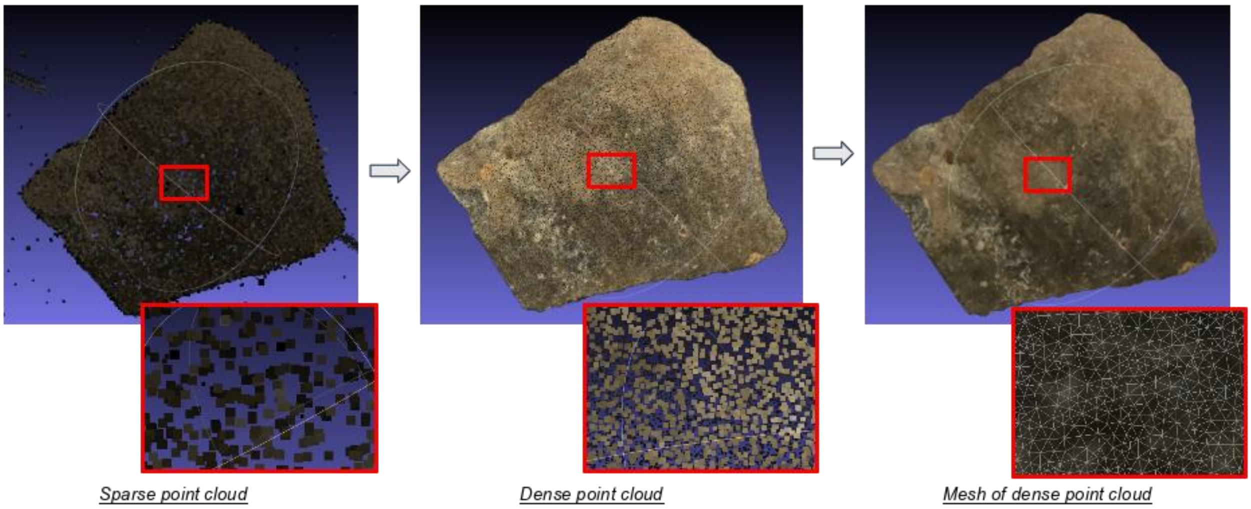

Figure 12 - Multi-view stereo pipeline of OpenMVS applied on the object Lot1_112

3.5 Limitations and future improvements

Although the proposed workflow provides an automated solution for the registration of hyperspectral images and the generation of detailed 3D models, several limitations remain. First, the complete processing sequence currently requires approximately one hour per artefact. This is largely due to the use of full-resolution images (8256x5504), which provide a high level of geometric detail but substantially increase computational cost. At this stage, an appropriate balance between reconstruction quality and processing speed has not yet been established. Sparse point cloud generation (SfM) can be performed incrementally during image acquisition, which can drastically reduce post-acquisition processing time. Second, the masking strategy relies on the quality of the sparse reconstruction generated by COLMAP. If an insufficient number of reliable features is detected during the SfM stage, the resulting sparse geometry may be incomplete, which can negatively affect both mask generation and the following dense reconstruction. Third, the performance of the Multi-View Stereo reconstruction remains dependent on the visual properties of the object surface. Highly reflective, transparent, or very dark areas may reduce the effectiveness of purely geometric approaches such as OpenMVS, as these methods rely on image intensity consistency and reliable pixel correspondences across multiple viewpoints.

Future developments could investigate learning-based approaches to improve specific stages of the workflow, particularly segmentation and dense reconstruction. For example, combining a learning-based segmentation method, such as BiRefNet (Zheng et al., 2024), with the geometric masking approach presented in this work could improve the robustness and consistency of object extraction across different viewpoints. Similarly, recent learning-based Multi-View Stereo methods could be evaluated for archaeological artefacts, whose surface properties often limit the performance of conventional photogrammetric reconstruction.

These complementary approaches will be assessed against the requirements of archaeological documentation and archaeometric analysis, including geometric reliability, repeatability, traceability, processing time and compatibility with the wider AUTOMATA workflow.

4 Reference Information System in 3D (RIS3D)

This section presents the RIS3D, its underlying concepts, and its main functionalities. The application can be downloaded and used to explore the enriched 3D models and associated multimodal data included in the D5.3 demonstration collection. This demonstration dataset will be progressively expanded as part of WP6, which aims to further automate the creation and visualisation of enriched 3D models. In subsequent acquisition campaigns, additional collections will be created, potentially with different access rights from those of this publicly available demonstration dataset. Section 4.1 describes the RIS3D architecture and the adaptations implemented within AUTOMATA, while Section 4.2 provides instructions for downloading, accessing and testing the application and its reference datasets.

4.1 RIS3D concepts and functionalities

The RIS3D (Dutailly et al., 2023) is a concept developed by Archeovision. It is closely related to a Geographic Information System (GIS) but operates entirely within a three-dimensional environment. It provides a common spatial framework in which heterogeneous information can be associated with points, areas or volumes defined in relation to a 3D reference model.

RIS3D is implemented as a software platform that links 3D reference datasets (or 3D models) with a heterogeneous and large-scale database. The system is designed to be generic and can be applied to any type of cultural heritage object, ranging from individual artefacts to large-scale archaeological sites.

Data stored in the database, such as physico-chemical analysis results, are spatially referenced using coordinates defined within the same coordinate system as the 3D reference datasets. However, these data remain independent of the 3D datasets, which can be replaced or exchanged in the same way that 2D raster layers are managed in a conventional GIS.

The information associated with each spatial entity is fully flexible and is stored in a hierarchical structure using key-value pairs. Keys are text labels used to identify individual pieces of information, while values may consist of text, numerical values, dates, images, files, 2D datasets, camera positions, chemical analyses (including qualitative or quantitative data: spectra, chemical composition, etc.), spatial datacube projected onto the 3D mesh (HSI or XRF mapping), or other data types.

The data are visualised using a dedicated 3D viewer capable of displaying the 3D reference datasets, spatial entities, and their associated information in real time. A layer management system enables users to filter and query the data. Queries can be performed on both the spatial entities and the associated metadata.

From a technical perspective, Archeovision provides an implementation based on a PostgreSQL database using the JSONB (JavaScript Object Notation Binary) storage format, a web server providing secure access with user rights management, and a 3D application developed in Unity for visualising both the 3D reference datasets and the associated database records. Within the AUTOMATA project, the RIS3D has been adapted for integration into the automated workflow developed for photogrammetric reconstruction and data integration.

More specifically, the software has been modified to correctly handle the files generated by the dedicated photogrammetric pipeline, as well as the camera positions calculated by COLMAP. These adaptations allow the spatial projection of hyperspectral images onto the 3D models. The way RIS3D processes and displays Raman and XRF spectra has also been adapted to support the specific file formats stored in the AUTOMATA database. These modifications enable the 3D geometry, hyperspectral projection and point-based archaeometric measurements to be explored within a spatial reference system.

In addition, the user interface has been simplified by removing functionalities that are not required within the scope of AUTOMATA. This optimisation reduces unnecessary complexity and makes the software better suited to the specific needs of the project workflow.

4.2 RIS3D implementation and use

This section provides practical instructions for accessing, installing and testing the AUTOMATA RIS3D application and the D5.3 demonstration collection. Unlike the preceding section, which describes the concepts, architecture, and project-specific adaptations of RIS3D, the following subsections adopt an operational format to guide users through software download, login, model loading, and visualisation of the associated HSI, pXRF, and Raman data.

The Windows build represents the currently validated release. A macOS build is also available, but it has not yet undergone equivalent testing.

4.2.1 Access and software download

The AUTOMATA RIS3D application can be downloaded at https://automata.3dris.science/.

> *For Windows systems, the available package is:* ***buildRIS3D - Automata.zip***
>
> *For macOS systems, the package* ***buildRIS3D - Automata - Mac.app.zip*** *is also available. However, this build has not yet been fully tested, and the Windows version should therefore be considered the reference release for D5.3.*

4.2.2 Installation and application launch

After download, the archive should be extracted and the application launched by running SIR3D.exe file from the extracted folder.

On some Windows systems, Microsoft Defender SmartScreen may display a warning because the application is not digitally signed. In this case, select the additional information option before choosing to run the application.

4.2.3 Access to the D5.3 demonstration collection

Log in using the username ‘demo’ and the password ‘demo’.

> *These credentials provide access to the public demo collection associated with D5.3. The collection contains the artefacts acquired using the fixed-object protocol developed to support automatic registration between the photogrammetric and hyperspectral datasets.*

RIS3D also supports the creation of separate collections with distinct access credentials. This functionality allows access to be managed according to the status, ownership and dissemination conditions of each dataset.

4.2.4 Downloading and loading 3D models

In the model panel shown in Figure 13, the Server tab provides access to the 3D models available in the selected collection. A model must first be downloaded and then loaded into the viewer. Download time depends on the model size and the available network connection.

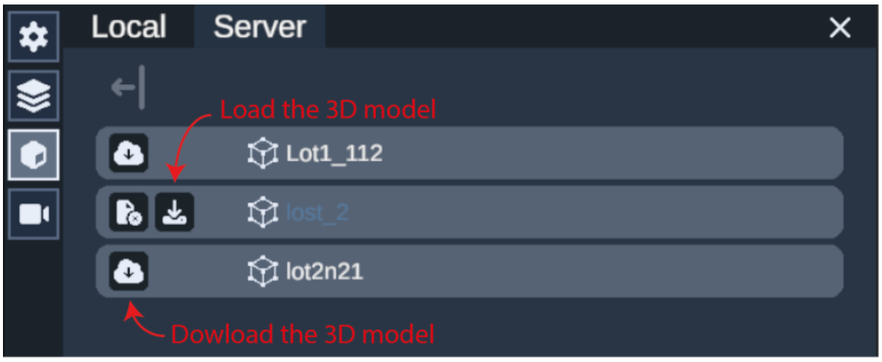

Figure 13 - RIS3D model panel used to download and load 3D models from the selected collection.

Once loaded, individual models can be shown  or hidden  through the corresponding visibility controls.

To make it easier to manipulate the object, the centre of rotation can be repositioned by holding Ctrl and left-clicking on the required point of the 3D model. The selected centre is indicated by a white marker. Since the marker retains a constant on-screen size, the view may need to be enlarged when working with small objects.

Lighting direction and intensity can be adjusted by pressing Alt + left-click or Alt + zoom.

Local 3D models can also be accessed through the Local tab. To make a model available in the application, the corresponding file must be placed in the following directory: buildRIS3D - Automata\buildRIS3D - Automata\SIR3D_Data\GLB

4.2.5 Selecting and loading analytical layers

In the  panel, 'Layers' tab, the information associated with the selected collection can be displayed  or hidden  using the visibility controls.

If a selected layer is not displayed correctly, the 'Reload' command  can be used to retrieve the associated data again.

The current implementation allows analytical layers to be selected independently of the loaded 3D model. For correct spatial visualisation, the selected layer and the loaded model should refer to the same artefact identifier. Future releases will further streamline this association within the user interface.

4.2.6 Visualising hyperspectral data

When an HSI layer is selected, a floating panel may appear in front of the 3D model with a 'Download' command. The same function can be accessed through the location icon shown in Figure 14. Selecting this command downloads the hyperspectral cube and initiates its projection onto the 3D model. Loading time depends on the size of the hyperspectral dataset and the available network connection.

Figure 14 - RIS3D Layers panel, including the control used to download and activate hyperspectral data.

In the HSI panel (Figure 15), select the model onto which you wish to project the image from the 'Show projection' drop-down menu.

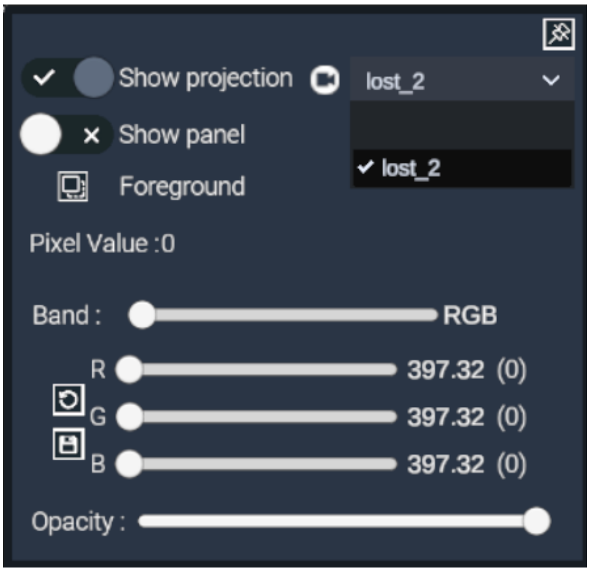

Figure 15 - HSI floating panel in RIS3D

Once a 3D model and its associated archaeometric layers have been selected, the enriched model can be explored as shown in Figure 16. The viewer supports switching between RGB and hyperspectral representations, selecting the wavelength displayed in the hyperspectral projection, and examining spatially referenced Raman and pXRF spectra.

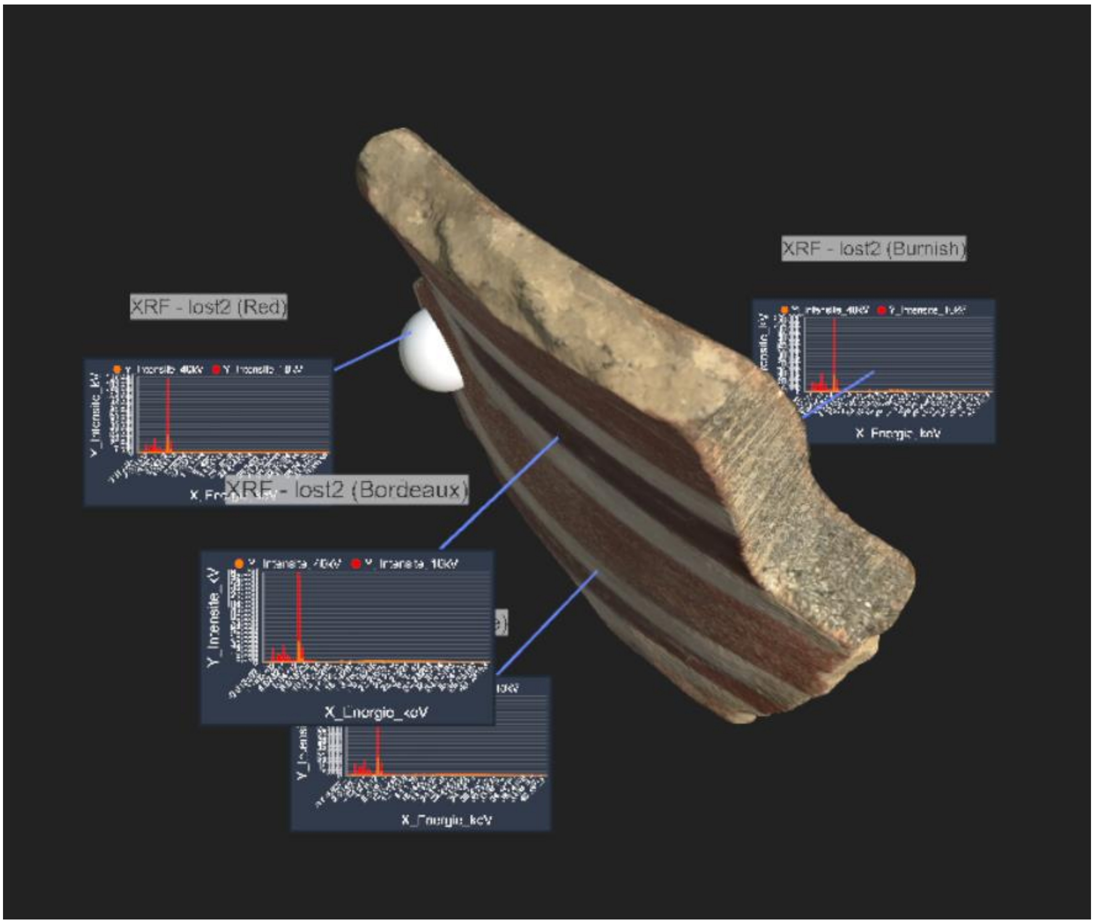

Figure 16 - Example of an enriched 3D model of a ceramic sherd as viewed in RIS3D, where four XRF analyses are represented by their spectra.

> *Raman and pXRF measurements may be associated with the same or closely adjacent locations on the 3D model. When both layers are displayed simultaneously, their visual markers or panels may overlap and partially obscure one another.*

Additional information can be accessed by holding Shift and left-clicking on an XRF or Raman spectra panel to display its metadata and view the spectra in a dedicated viewer (Figure 17) by clicking on .

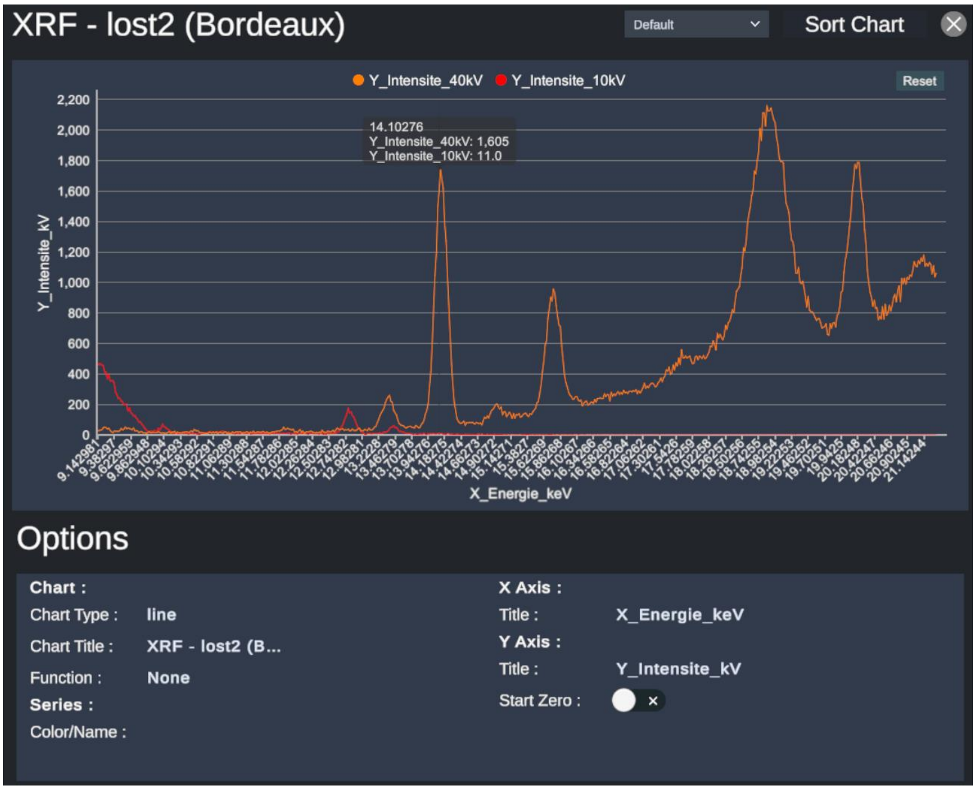

Figure 17 - Example of an XRF spectrum in the dedicated viewer of RIS3D with a zoom on the strontium element X-Ray lines.

5 Automatic detection of Regions of Interest (ROIs)

An automated workflow was developed to identify and rank ROIs for pXRF and Raman spectroscopy. The method combines hyperspectral information with the geometric properties of the photogrammetric model. Hyperspectral imaging is used to distinguish spectrally different surface regions, while the 3D geometry is used to identify locations that are visible, reachable, sufficiently flat, and safely separated from rims, fractures, and other geometrically unsuitable areas. For each artefact, the workflow receives three principal inputs:

1. the photogrammetric surface mesh (in .glb format);
2. the hyperspectral imaging (HSI) cube in ENVI format, consisting of a RAW or DAT file and the corresponding HDR header;
3. the HSI viewfinder or ROI images associated with the HSI acquisition.

When available, the reflectance-corrected hyperspectral cube produced by the camera processing software is used in preference to the uncorrected radiance data.

5.1 Hyperspectral segmentation of surface regions

The first stage identifies spectrally distinct regions on the visible surface of the artefact. These regions represent candidate material or surface classes, such as the ceramic body, slip, decoration, deposits, encrustations or altered areas. The segmentation does not, by itself, constitute a definitive chemical identification, but provides a data-driven subdivision of the surface that can subsequently be investigated and validated through Raman or pXRF measurements.

The artefact is initially separated from the acquisition background using the BiRefNet image-segmentation model (Zheng et al., 2024). Since BiRefNet was trained on RGB images rather than hyperspectral data, the segmentation is performed on the HSI viewfinder image (RGB). Alternatively, if an artefact mask (ROI) has already been generated during the photogrammetric process, it can be provided directly, avoiding the need for automatic segmentation. To improve the robustness of the segmentation, the image is currently cropped to a manually selected bounding box extending approximately 5 cm around the artefact. This allows the model to focus on the object while reducing the influence of the surrounding scene and has consistently produced successful segmentations for all artefacts processed so far. No additional preprocessing is applied before inference, and BiRefNet is used with its default configuration. In future developments, this bounding box will be generated automatically from the object position estimated by the robotic acquisition system, allowing the artefact to be centred without manual intervention. The resulting binary segmentation mask is then applied to the hyperspectral cube, restricting the subsequent spectral analysis to pixels belonging to the artefact while removing the influence of the support, acquisition table and surrounding scene.

Each retained spectrum is then processed using standard normal variate (SNV) normalisation, i.e. for each pixel, the mean spectral intensity is subtracted, and the result is divided by the spectral standard deviation. This operation reduces multiplicative and additive intensity differences caused by illumination, surface orientation and overall reflectance, while preserving the characteristic shape of the spectral signature. A light spatial smoothing operation is subsequently applied to suppress isolated pixel-level noise without removing the principal spatial boundaries between surface regions. Principal component analysis (PCA) is used to reduce the dimensionality of the hyperspectral cube. The number of retained components is selected so that approximately 99% of the spectral variance is preserved. This transformation reduces computational cost and limits the effect of redundant or noisy bands while retaining the spectral information required for clustering.

The reduced spectral representation is clustered using Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN, Campello et al., 2015; McInnes et al., 2017). HDBSCAN was selected because it does not require the number of clusters to be specified a priori, can identify clusters with varying densities, and explicitly distinguishes points that do not belong to any cluster as noise. These characteristics make it particularly suitable for hyperspectral data generated from archaeological material, where the number, size, and distribution of spectrally distinct surface regions is generally unknown, but is certain to vary between artefacts. The cluster assignments are subsequently mapped back onto the original two-dimensional image. Pixels previously excluded because of specular reflections remain unlabeled, while samples classified as noise by HDBSCAN are initially retained as noise. Several post-processing operations are then applied to improve the spatial coherence and practical usability of the segmentation. Clusters containing fewer than 20 pixels are discarded, and their pixels are transferred to the unassigned pool. In addition, an edge band extending approximately 10 pixels from the boundary of the artefact mask is computed. If more than 50% of the pixels belonging to a cluster fall within this band, the entire cluster is reclassified as noise, since it is assumed to correspond to background or shadow leakage introduced at the mask boundary rather than a genuine material region. All unassigned pixels, including HDBSCAN noise, specular-reflection pixels, discarded small clusters, and edge-rejected clusters, are subsequently evaluated together. When they comprise at least 200 pixels, they are grouped into a single additional cluster; otherwise, they remain labelled as noise. This step allows potentially valid surface regions that were fragmented or rejected during density-based clustering to be retained for subsequent analysis. Finally, the remaining valid clusters are reassigned consecutive identifiers and visualisation colours, while pixels labelled as noise remain unlabeled so that they are ignored during the subsequent automated target-selection workflow. The final result is a labelled image in which each retained cluster approximates a distinct spectral material or surface type.

5.2 Registration of the hyperspectral image and the 3D model

The registration procedure described here is complementary to the photogrammetric registration presented in Section 3 and addresses a different objective. This geometry-driven formulation is particularly suited to the automated ROI-selection workflow, where only the segmented HSI image and the reconstructed mesh are required to establish the correspondence between spectral regions and the 3D surface.

The HSI acquisition represents the upward-facing side of the artefact. The registration stage estimates the corresponding orientation of the 3D mesh and establishes a transformation between hyperspectral image coordinates and the rendered mesh.

Geometric descriptors are first computed for the mesh vertices. These include surface normals, curvature evaluated at both local and coarser spatial scales, and a concavity-depth measure expressing the distance of each point from the convex hull of the object. The descriptors are subsequently used to evaluate surface flatness and to penalise locations situated within deep concavities.

A set of candidate camera viewpoints is generated around the mesh. The principal axes derived from PCA of the mesh geometry are used to prioritise the dominant faces of the artefact. This is particularly suitable for approximately planar objects that generally present two principal opposing surfaces. For each candidate viewpoint, the mesh silhouette is rendered and compared with the object outline extracted from the hyperspectral cluster map.

The comparison is repeated over a range of in-plane rotations. Registration quality is evaluated using the intersection-over-union metric between the HSI object mask and each rendered silhouette. The viewpoint and rotation producing the highest overlap are retained. This procedure provides an estimated camera pose and a two-dimensional transformation between HSI pixels and the rendered representation of the mesh.

5.3 Visible-surface mapping and exclusion of unsafe areas

Following registration, the analysis is restricted to the surface visible from the recovered HSI camera position. A regular grid of rays is cast from the virtual camera through the image plane and onto the mesh. For each ray, the first intersected triangle and its corresponding 3D position are recorded. The resulting lookup structure, referred to as the face buffer, establishes a direct relationship between HSI pixels and locations on the mesh. Triangles that are never intersected are considered part of the hidden side of the artefact and are excluded from subsequent target selection. This visibility constraint prevents an ROI from being assigned to the back surface or to a geometrically occluded area.

Following an initial evaluation by archaeometric specialists within the AUTOMATA consortium, additional geometric constraints were introduced to better reflect practical acquisition requirements. In particular, locations situated close to artefact edges, fracture surfaces, or rounded rims were identified as problematic for acquisition. To incorporate this insight, a separate rim and steep-surface map is generated to identify potentially unsafe locations. Vertices whose smoothed surface normal deviates by more than approximately 55° from the viewing direction are classified as steep. These vertices generally correspond to the outer rim, fracture surfaces, ridge flanks or abrupt topographic transitions. Their positions are stored in a nearest-neighbour search structure, allowing the distance between any candidate ROI and the closest steep region to be evaluated efficiently.

Only camera-visible surfaces contribute to this rim model. Excluding the hidden side is important for thin artefacts, since otherwise the proximity of the opposite surface could cause most of the visible face to be incorrectly classified as close to an edge.

5.4 Selection of pXRF and Raman targets

Candidate pXRF locations are generated directly from the vertices belonging to camera-visible mesh faces. Where necessary, the candidate set is spatially subsampled to maintain computational efficiency while preserving uniform coverage of the visible surface.

For each candidate vertex, the algorithm attempts to fit a locally planar square representing the pXRF measurement footprint. The square is constructed in the tangent plane defined by the vertex normal. The search begins with the nominal instrument footprint, corresponding approximately to a 1 cm target area when the model scale is reliable. Progressively smaller footprints are tested when the nominal size is incompatible with the mesh dimensions or when no valid location can otherwise be identified. All vertices contained within the candidate footprint are evaluated. The location is retained only when their normals agree with the central normal within an angular tolerance of approximately 18°. This criterion rejects highly curved, irregular or discontinuous surface patches on which stable probe placement or reliable measurement geometry would be difficult to achieve.

Using the rim and steep-surface map, locations close to potentially unsafe locations are penalised in the ranking process. Although such areas may satisfy the local flatness criterion, they often correspond to thin or geometrically unstable regions where reliable pXRF measurements are more difficult to obtain and instrument positioning may be compromised. Consequently, the ranking function favours areas away from the edges whenever suitable alternatives are available, while still allowing near-edge locations to be selected as lower-priority targets when no better candidates exist. Initial assessment indicates that this refinement improves the agreement between the automatically selected targets and those typically chosen by an experienced operator.

Each valid candidate is ranked using a weighted objective function combining local surface roughness, variation in surface-normal orientation and concavity depth. Surface flatness is the principal selection criterion. The ranking also incorporates penalties for proximity to artefact edges, fracture surfaces and rounded rims, reflecting practical constraints identified during validation of the workflow. A graded penalty is additionally applied according to the distance from rims and steep surfaces. Candidates located at least approximately 1 cm from a steep region receive no rim penalty. At shorter distances, the penalty increases progressively and may reach approximately twice the full flatness-score range. This formulation strongly favours safe areas away from the edges while still allowing a near-edge target to be selected when no more suitable location exists.

The three highest-ranking non-overlapping locations are retained as pXRF targets. A diagnostic image displays the selected locations over a greyscale flatness map. The principal outputs of this stage are:

- the selected targets on the 3D model;
- a diagnostic flatness and target-selection image;
- a JSON report containing the target ranks, 3D coordinates, geometric scores and associated processing parameters.

Raman targets are selected separately for each hyperspectral cluster so that the principal spectrally distinct surface regions are represented. The cluster masks are extracted from the labelled PNG image. Colour values are quantised before extraction, and very small connected components are removed as probable segmentation noise.

The Raman workflow first attempts to reuse the pXRF targets. Each selected pXRF coordinate is projected into HSI image space using the recovered registration. When the projected location falls within a valid spectral cluster, it is adopted as the first Raman target for that cluster. This reduces the total number of required acquisition positions and ensures that, where appropriate, elemental and molecular measurements are obtained from the same surface area. The same geometric considerations adopted for pXRF target selection are applied to the Raman workflow, favouring stable and easily accessible surface regions while avoiding artefact edges and rounded rims whenever possible. Additional targets are selected until a maximum of three Raman locations per cluster is reached. Candidate positions are generated from the connected components of each cluster mask. When a cluster consists of a single extended component, the region may be subdivided along its principal axis to distribute the targets across the full spatial extent of the material rather than concentrating them in one area. An interior-distance filter is used to retain pixels sufficiently far from the cluster boundary. Each candidate pixel is mapped to a 3D location through the face buffer and is accepted only when the intersected mesh face is consistent with the corresponding cluster region. This verification prevents targets from being transferred across boundaries as a consequence of minor registration or rasterisation errors.

The candidate score combines normalised local flatness, a moderate preference for the centre of the connected component, the graded rim-proximity penalty, and a penalty for proximity to neighbouring spectral clusters. A minimum separation is also imposed relative to all previously selected pXRF and Raman targets. The penalties are continuous rather than binary. Consequently, when a spectral cluster is located predominantly on a rim or geometrically difficult surface, the algorithm identifies its safest and flattest available position instead of automatically rejecting the complete region. Clusters for which no geometrically viable target can be identified are omitted and recorded in the processing report. This prevents the generation of coordinates that cannot be reached safely or that do not provide a sufficiently stable analytical surface. The Raman results are exported as:

- the coloured spectral regions and plus-shaped Raman crosshairs;
- a regions-only file containing the projected cluster classification;
- a diagnostic PNG showing the registration, candidate regions and selected targets;
- a JSON report containing the cluster identifiers, target ranks, 2D and 3D coordinates, target origin and geometric quality indicators.

The complete workflow therefore produces ranked analytical targets that are simultaneously informed by hyperspectral variability, surface geometry, camera visibility and instrumental placement constraints. These outputs provide the spatial information required by the automated acquisition system to position the pXRF and Raman sensors and to associate each measurement with its corresponding location on the enriched 3D model.

The proposed pipeline was applied to a dataset of 33 archaeological artefacts, covering the range of object morphologies and surface characteristics available within the AUTOMATA project. At present, the only manual input required is the specification of the object's bounding box. This dependency is expected to be removed in the complete AUTOMATA system, where object localisation and bounding-box estimation will be provided automatically by the robotic acquisition framework.

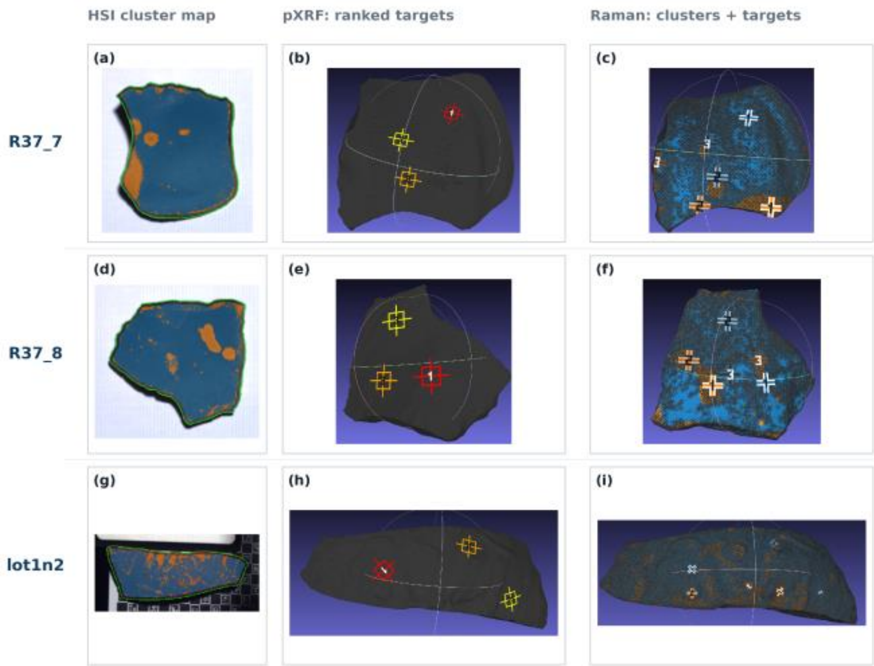

Figure 18 - Automated workflow for the detection and selection of ROIs on two lithics (R37_7 and R37_8) and one ceramic artefact (lot1n2). Hyperspectral segmentation and spectral clustering produce labelled surface-region maps (left). These maps are registered with the corresponding photogrammetric models, and the analysis is restricted to camera-visible surfaces while geometrically unsafe areas are penalised or excluded. The resulting spatial information supports the selection of the 3 highest-ranked pXRF targets (centre) and cluster-informed Raman targets (right).

5.5 Code availability

The implementation of the methods presented in this section is publicly available as open-source code on GitHub at the link https://github.com/minstudio/automata_crosshair. The repository contains the implementation of the workflow described in Sections 5.1 to 5.4, including hyperspectral segmentation of surface regions, visible-surface mapping and exclusion of unsuitable acquisition areas, and the automatic selection of ROIs for pXRF and Raman analyses. The published code is intended to support the reproducibility of the proposed methodology and to facilitate its integration and further development within the AUTOMATA automated acquisition pipeline.

6 Conclusions and future development

This deliverable presents the developments carried out within WP5 to improve the acquisition, processing, and integration of multimodal datasets for enriched 3D digitisation. It follows T5.2 (*Creation of 3D models for pottery and for lithics*) with a dedicated RIS3D server database which is available online (https://automata.3dris.science/). Building upon the metadata model and database infrastructure established in D5.1 and D5.2, the work focused on establishing a consistent workflow for the enrichment of archaeometric data within the photogrammetric 3D reconstruction of artefacts, providing a common spatial reference for future visualisation and analysis of cultural heritage objects. The resulting methodologies and software developments constitute the final technological components required to complete the objectives of WP5.

Following the work on D5.2 *3D Database*, a slightly modified and improved acquisition setup was designed to ensure that hyperspectral and photogrammetric data are captured under identical conditions without moving the object, improving the robustness of the registration process. From these acquisitions, a reconstruction of the hyperspectral data cube and the photogrammetric images were integrated into the photogrammetric pipeline to register the hyperspectral camera at the same time as 3D reconstruction. With the developments realised for automatic detection of Regions of Interest, the ROIs are automatically chosen for Raman and XRF analysis, which will further be integrated into the RIS3D. Beyond the technical developments, the acquisition campaigns carried out validated the proposed methodology under realistic operating conditions. The resulting datasets constitute both reference material for evaluating the complete processing chain and practical test cases for the continued development of the AUTOMATA platform. They also contributed to identifying acquisition constraints, processing requirements, and integration challenges that will guide the next implementation phases of the project.

The developments presented in this deliverable establish the methodological and technical foundations for the activities of WP6 *Technologies for visualisation and processing of enriched digitisation*. Future developments will focus on the complete automation and integration of the data processing workflow. This will include direct ingestion and automated processing of data collected from all acquisition devices, the transfer of these datasets to the photogrammetric reconstruction workflow, and their integration into the project database, as planned in D6.1 *Direct Ingesting*. Further developments will also address the implementation of online visualisation capabilities within the RIS3D platform for D6.2 *Visualisation and Interaction*, enabling users to explore enriched 3D models directly through a web interface. The current implementation of RIS3D, together with the publicly available demonstration dataset, will also provide an opportunity to gather feedback from users. This feedback will be used to refine the platform, improve its usability and ergonomics, and, where necessary, adapt or extend its functionalities to better meet users' needs. Finally, communication mechanisms between the RIS3D acquisition system and the Archeogrid infrastructure will be established to ensure efficient data exchange and support a fully integrated processing workflow across the different components of the project.

Taken together, these developments represent an important milestone towards the fully automated enriched digitisation pipeline described by the AUTOMATA project. By integrating acquisition, processing, registration and data management within a unified framework, the project establishes the technological basis required for the large-scale production and exploitation of enriched digital twins of cultural heritage objects.

References

Bleyer, M., Rhemann, C., & Rother, C. (2011). PatchMatch stereo: Stereo matching with slanted support windows. In BMVC, 11, 1–11.

Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. (2015). Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection. ACM Transactions on Knowledge Discovery from Data, 10(1), 1–51

Cernea, D. (2020). OpenMVS: Multi-View Stereo Reconstruction Library (Version 2.4.0) [GitHub Repository]. https://cdcseacave.github.io

Dutailly, B., Portais, J.-C., & Granier, X. (2023). RIS3D: A Referenced Information System in 3D. Journal of Computing and Cultural Heritage, 15, 73:1–73:20. https://doi.org/10.1145/3517043

Fei, B., Xu, J., Zhang, R., Zhou, Q., Yang, W., & He, Y. (2024). 3D Gaussian splatting as new era: A survey. IEEE Transactions on Visualization and Computer Graphics.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In Proceedings of the Seventh IEEE International Conference on Computer Vision, 2, 1150–1157.

Maté-González, M. Á., Yali, R., Rodríguez-Hernández, J., González-González, E., & Aguirre de Mata, J. (2025). Comparison of NeRF- and SfM-based methods for point cloud reconstruction for small-sized archaeological artifacts. Remote Sensing, 17(14), 2535.

Matl, M. (2019). Pyrender (Version 0.1.45) [GitHub Repository]. https://github.com/mmatl/pyrender

McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. Journal of Open Source Software, 2(11), 205. https://doi.org/10.21105/joss.00205

Schönberger, J. L., & Frahm, J. M. (2016). Structure-from-Motion Revisited. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4104–4113.

Umeyama, S. (1991). Least-squares estimation of transformation parameters between two point patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(4), 376–380.

Wang, F., Zhu, Q., Chang, D., Gao, Q., Han, J., Zhang, T., ... & Pollefeys, M. (2026). Learning-based multi-view stereo: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 48, 6241–6260.

Yao, M., Huo, Y., Ran, Y., Tian, Q., Wang, R., & Wang, H. (2026). Neural radiance field-based visual rendering: A comprehensive review. IEEE Transactions on Visualization and Computer Graphics.

Zheng, P., Gao, D., Fan, D. P., Liu, L., Laaksonen, J., Ouyang, W., & Sebe, N. (2024). Bilateral Reference for High-Resolution Dichotomous Image Segmentation. CAAI Artificial Intelligence Research, 3, 9150038.