RESEARCH ARTICLE

Open Access

Image-based metric heritage modeling in the near-infrared spectrum

Efstathios Adamopoulos[1*], Alessandro Bovero[2] and Fulvio Rinaudo[3]

Abstract

Digital photogrammetry and spectral imaging are widely used in heritage sciences towards the comprehensive recording, understanding, and protection of historical artifacts and artworks. The availability of consumer-grade modified cameras for spectral acquisition, as an alternative to expensive multispectral sensors and multi-sensor apparatuses, along with semi-automatic software implementations of Structure-from-Motion (SfM) and Multiple-View-Stereo (MVS) algorithms, has made more feasible than ever the combination of those techniques. In the research presented here, the authors assess image-based modeling from near-infrared (NIR) imagery acquired with modified consumer-grade cameras, with applications on tangible heritage. Three-dimensional (3D) meshes, textured with the non-visible data, are produced and evaluated. Specifically, metric evaluations are conducted through extensive comparisons with models produced with image-based modeling from visible (VIS) imagery and with structured light scanning, to check the accuracy of results. Furthermore, the authors observe and discuss, how the implemented NIR modeling approach, affects the surface of the reconstructed models, and may counteract specific problems which arise from lighting conditions during VIS acquisition. The radiometric properties of the produced results are evaluated, in comparison to the respective results in the visible spectrum, on the capacity to enhance observation towards the characterization of the surface and under-surface state of preservation, and consequently, to support conservation interventions.

Keywords: Image-based modeling, Photogrammetry, Near-infrared imaging, Modified cameras, Cultural heritage survey

Introduction

Multi-view image recording

In the course of the last decade, rapid advancements in passive sensors for 3D recording, workflows for swift data acquisition, automatic or semi-automatic software which implement image-based reconstruction algorithmic approaches and computational systems for processing of large datasets, have taken place. As a result, the use of 3D image-based modeling technologies has become common for various aspects of heritage science.

Multi-view image recording [1, 2] has prevailed as a low-cost, efficient, and easily implementable technique for cultural heritage documentation, interpretation, dissemination, and protection [3–5]. Multi-view image approaches facilitate the digitization of tangible heritage with reduced needs for supervision and expertise. At the same time, they enable the production of accurate and high-resolution results with images from relatively low-cost digital cameras. These approaches differ from traditional close-range photogrammetry, due to the possibility to use oblique imagery and to simultaneously estimate the internal and external orientation camera parameters, without the need for a prior definition of control points of known coordinates on a reference system [6, 7]. However, the implementation of control points is recommended during orientation for more accurate results and is mandatory for georeferencing. For applications that only require scaling in a local coordinate system, a simple, measured distance can be adequate. The processing

*Correspondence: efstathios.adamopoulos@unito.it
[1] Department of Computer Science, Università degli Studi di Torino, Corso Svizzera 185, 10149 Turin, TO, Italy
Full list of author information is available at the end of the article

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 2 of 12

pipeline starts with feature detection and description on every image of the dataset. Then follows a Structure-from-Motion (SfM) implementation to determine camera positions in 3D space and the coordinates of the scene, producing a sparse point cloud. Given the camera orientations, dense image matching algorithms enable further densification of the point cloud, as almost every pixel of the scene is reconstructed in 3D -a procedure typically named Multiple-View-Stereo (MVS) or dense stereo-matching. Later, these dense point clouds can be transformed into textured models via surfacing algorithms and texture mapping.

Applications of terrestrial multi-view image-based modeling, and texturing, for heritage applications, vary considerably [8, 9]. This technique has been implemented towards high-resolution recording and evaluation of the state of preservation of stone [10], wood [11], and painted artworks [12]. Although, unless measurable geometrical changes have occurred on the object, close-range photogrammetry with visible imagery has little to offer for the assessment of degradation on its own.

Near-infrared imaging and modeling

In the context of many heritage analyses, and especially for the study of polychromatic artworks and the investigation of surface or undersurface deterioration of historical artifacts, the use of visible-spectrum textures is often not adequate. Near-infrared imaging has often been explored towards this direction [13, 14], with sensors that employ complementary metal-oxide-semiconductors (CMOS) based on InGaAs (indium gallium arsenide, 750–1700 nm) or PtSi (platinum silicide, 750–5000 nm) detectors, developed in the 1990s [15–17].

Recently, the use of commercial digital cameras employing CCD (charge-coupled device) and CMOS sensors, modified for near-infrared or full-spectrum acquisition (combined with external NIR filters) has been commonized in heritage science. They consist of lower cost, high-resolution alternatives with spectral imaging capabilities, which retain user-friendly features and interfaces to a broad variety of photographic accessories and software [18–22].

The availability of high-resolution, easily operated digital cameras for NIR acquisition, in combination with SfM/MVS image-based modeling techniques, has made feasible the 3D spectral modeling for heritage applications. Contemporary research, which showed promising results [22–28], led us to the motivation for further experimentation.

Research aims

In this work, we evaluate the use of imagery from consumer-grade digital reflex cameras modified for near-infrared imaging, in combination with image-based 3D reconstruction techniques, to produce high-fidelity models of tangible heritage assets textured with spectral information. To perform a thorough evaluation of this combined approach, we have acquired rigid datasets of images for case studies of archaeological importance with varied geometry and characteristics, and maintained constant the parameters of the acquisition for the visible and non-visible spectra involved. We produced 3D meshes utilizing software of different algorithmic implementations and compared the data produced by each method, in terms of surface deviation. Furthermore, we explored the usability of produced spectral textures. The principal aims of our research were to objectively quantify the quality of the 3D models produced by this integrated spectral modeling method, and to assess its applicability for different heritage case studies. In addition, we attempted to understand how this method can potentially enhance the preservation of surface detail on the reconstructed 3D digital models. Finally, we aimed to evaluate the capacity of the fabricated near-infrared textures, for augmenting observations towards the characterization of the state-of-preservation.

Materials and methods
Case studies

To evaluate the presented approach, which combined multi-view photogrammetric modeling and near-infrared spectral imaging (from modified consumer cameras), we applied it for four historical artwork case studies. The objects to be studied were chosen due to their dissimilar dimensions, shapes, surface roughness, reflectivity and environmental conditions of preservation (Fig. 1). The first one was a 17th century marble statue (approx. height 1.75 m; approx. base dimensions 0.45×0.35 m^2) from the Fountain of Hercules at the Reggia di Venaria Reale near Turin. The Palace of Venaria was one of the Residences of the Royal House of Savoy, included in the UNESCO World Heritage List in 1997. The statue is owned by Musei Reali Torino.

The second case study was a 18th century Chinese Coromandel folding screen (2.67×1.94 m^2) from Castello di Grinzane Cavour (Santena, Province of Turin, Piedmont). The castle's history starts around 1350, it was owned by several lords, the most famous of which was, in the mid-19th century, Camillo Benso, Conte di Cavour, who resided there starting from 1830 and restored the construction. It was inscribed on UNESCO World Heritage List in 2014. The object is owned by the City of Turin–Cavour Foundation.

The third case study was an 18th century wooden furniture part painted with flowers (2.39×0.81 m^2), from the Sala dell'Alcova in Palazzo Chiablese. The Palazzo

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 3 of 12

Fig. 1 Captured VIS images of the case studies (from left to right) statue from the Fountain of Hercules in Venaria (Turin), Coromandel folding screen, wooden furniture part decorated with flowers, small sculpture of Crist Crucified

Chiablese is a wing of the Royal Palace of Turin. It was the residence of the Duke of Chablais first and then of Carlo Felice, King of Sardinia, and Ferdinando, Duke of Genoa. The object is owned by the Italian Ministry of Cultural Heritage and Activities.

The fourth, and last, case study was a small 19th century religious stone sculpture of Christ crucified (0.31×0.22 m^2) from Castello di Casotto (Garessio, Province of Cuneo, Piedmont) owned by the Region of Piedmont. The Casotto Castle was originally a Carthusian monastery, later acquired by the Savoy and transformed into a castle and hunting lodge by Carlo Alberto.

Instrumentation, software and tools

NIR images were collected using a Nikon D810 professional camera modified for near-infrared acquisition (1650$) and a low-cost Canon EOS Rebel SL1 camera modified for full-spectrum acquisition (500$). The D810 is a Digital Single-Lens Reflex (DSLR) camera that employs a full-frame CMOS sensor (35.9×24 mm^2, 4.88 μm pixel size), with a maximum resolution of 36.3 effective Megapixels. It was modified by replacing the internal NIR-cut filter (that would not allow acquisition in the near-infrared spectrum) with an internal NIR-pass filter by the company MaxMax LDP, which essentially transforms the camera into an imaging sensor sensitive only in the near-infrared spectrum. The Rebel SL1 is a DSLR camera that employs an APS-C CMOS sensor (22.3×14.9 mm^2, 4.38 μm pixel size), with a maximum resolution of 18.0 effective Megapixels. It was modified by removing the internal NIR-cut filter by the company LifePixel, thus making the camera sensitive to a spectrum wider than the visible,

approximately between 280 nm and 1400 nm. For the Rebel SL1 camera, we used an external NIR-pass (700–1400 nm) filter to capture the near-infrared images. VIS images were captured using a Canon EOS 1200D compact DSLR that employs an APS-C CMOS sensor (22.3×14.9 mm^2, 4.40 μm pixel size), with a maximum resolution of 18.0 effective Megapixels. Also, the modified Canon EOS Rebel SL1, described above, was additionally used in combination with an external VIS-pass filter. For the Nikon camera, a 24 mm prime lens (160$) was used. For the Canon cameras, an 18–55 mm zoom lens (50$) was used. In order to avoid camera microshake effects and to produce better quality photos, the cameras were mounted on a tripod. For all photo-editing operations, we used Adobe Lightroom Classic.

Multi-view image-based reconstruction was conducted with Agisoft Metashape Professional 1.5.1 (AMP) for the first case study, and additionally with 3DFlow Zephyr Aerial 4.519 (FZA) [29] for the rest of the case studies, in order to also compare the performance of different algorithmic implementations with NIR imagery. Both commercial software are SfM/MVS approach based. Agisoft Metashape Professional employs SIFT-like detection and description, then calculates approximate camera location and used Global bundle-adjustment to refine them, a type of MVS disparity calculation for dense reconstruction and Screened Poisson surface reconstruction. 3DFlow Zephyr Aerial employs a modified Difference-of-Gaussian (DoG) detector, a combination of Approximate Nearest Neighbor Searching, M-estimator Sample Consensus and Geometric Robust Information Criterion for matching. Then uses hierarchical SfM and Incremental

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 4 of 12

adjustment, and dense MVS reconstruction with fast visibility integration, tight disparity bounding, Finally, in FZA, meshing with an edge-preserving algorithmic approach was selected, to differentiate from AMP.

Data acquisition

For the first two case studies, lighting conditions were not optimal to acquire photogrammetric datasets. The marble statue was illuminated non-uniformly by various spotlights. Thus, it was decided not to use flash as it resulted in more shadows at the occluded areas instead of reducing them. The Coromandel lacquerware was also under non-uniform lighting, with glaring effect visible on optical photos at the upper part, and flash could not be used, because of the highly reflective nature of the lacquer. For the rest of the case studies, we were able to appropriately utilize the built-in flash of the cameras, thus eliminating shadows, to acquire imagery datasets with as homogeneous radiometric characteristics as possible. An x·rite ColorChecker® Classic with 24 colors was used for color balancing, utilizing middle gray for the visible-spectrum datasets. Scaling was performed with an invar scale bar of 1.000165 m ($\pm$ 15 nm), barring for the small stone sculpture, where we used as reference the dimensions of the wooden cross. Additionally, small targets were placed on the base and body of the marble statue to facilitate further metric comparisons.

Dense acquisition of images was planned in such a manner as to acquire rigid datasets with large overlaps (>80%) from a very close range. Furthermore, we maintained the capturing conditions (internal parameters, external orientation parameters, and lighting) constant between VIS and NIR spectra, taking into consideration the different sensors employed and the surrounding conditions to obtain comparable results. For the first case study, we were not able to use exactly the same camera positions for VIS and NIR imaging, but we marked and maintained a constant distance from the object in both cases and approximately the same angles between each

position. For case study 4, we used a turntable during the acquisition, and we rotated the object four times after placing different sides towards the camera. Very low ISO values were used to prevent sensor luminance noise, simultaneously maintaining exposure under the clipping limit value. All images were acquired in RAW format. Acquisition conditions are summarized in Table 1. The Ground-Sampling Distances (GSDs) that we present in Table 1 refer to an approximate average of the size of the pixel on each object, which was calculated inside the SfM/MVS based software.

A STONEX F6 SR structured light scanner was used to perform control surveys of the case studies, thus creating reference models of the objects to perform metric comparisons. The F6 SR scanner has an accuracy of 90 μm at 250 mm distance, an effective range of 250–500 mm and a resolution of 0.4 mm at 250 mm distance. All models produced by scanning were down-sampled to match the density of the photogrammetric models.

Data processing

Image-based reconstruction from visible and near-infrared imagery followed a semi-automatic SfM/MVS workflow, standard for large-scale archaeological photogrammetric applications [30, 31], very similar on both employed software, despite considerable differences on the algorithmic implementations. The implemented images were masked accordingly, to exclude the unwanted areas of each scene. In the last case study, out-of-focus (blurry) areas on the stone sculpture, were also masked, to increase the quality of imagery, thus reducing noise levels and processing times. Between each stage of the reconstruction, thorough visual checks were performed to determine quality. Then, denoising procedures were followed, in an identical manner for all the produced meshes, to not reduce the comparability of results.

The 3D models were semi-automatically generated in a four-step process. The first step included a sparse reconstruction of each scene, with a simultaneous

Table 1 Characteristics of acquired image datasets

Case study	Camera model	MP	f	Pixel size	Distance	GSD [mm]	Spectrum	No. images	f-stop	Exposure	ISO
1	EOS 1200D	18.0	18.0 mm	4.40 μm	0.98 m	0.2 mm	VIS	180	f/11	1/2 s	200
	D810	36.3	24.0 mm	4.88 μm	1.24 m	0.2 mm	NIR	180	f/11	5 s	200
2	REBEL-SL1	18.0	18.0 mm	4.38 μm	2.57 m	0.6 mm	VIS	22	f/16	1 s	100
	REBEL-SL1	18.0	18.0 mm	4.38 μm	2.68 m	0.6 mm	NIR	22	f/16	1 s	100
3	REBEL-SL1	18.0	18.0 mm	4.38 μm	0.72 m	0.16 mm	VIS	40	f/16	2.5 s	100
	REBEL-SL1	18.0	18.0 mm	4.38 μm	0.72 m	0.16 mm	NIR	40	f/16	2.5 s	100
4	REBEL-SL1	18.0	55.0 mm	4.38 μm	0.26 m	0.02 mm	VIS	60	f/16	1/15 s	100
	REBEL-SL1	18.0	55.0 mm	4.38 μm	0.26 m	0.02 mm	NIR	60	f/16	1/15 s	100

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 5 of 12

approximate calculation of cameras' relative orientation at the moment of image acquisition, and autocalibration, with SfM approaches. For this step, the selected accuracy and density parameters were the highest available in both software. The sparse point clouds were cleaned according to reprojection errors, and local cluster distances with statistical filtering. During the second step, the results were densified by employing MVS stereo-matching algorithms. The third step consisted of meshing the dense 3D point clouds into triangular surfaces (3D Delaunay algorithm). The produced meshes were then cleaned from small unconnected components and spikes. The final step referred to the application of texture mapping to obtain single-file high-resolution textures from the original photographs. Given the high quality of the original imagery, we limited color balancing and blending between images to reduce the possibility of radiometric errors. The processing parameters inside both software were selected to minimize surface noise and to optimize the final textured results. However, at the same time, we attempted to maintain as similar parameter values as possible between software and spectra, to ascertain the metrological validity of the conducted research. Into choosing the resolution of textures, we considered sampling distances to be at least two or three times higher than the original pixel sizes.

Results and discussion

With the intention of evaluating the photogrammetric results we recorded in detail the processing times, reconstruction errors and volumes of the results, catalogued in Table 2. Detailed geometric comparisons were performed, to assess surface deviations among visible and near-infrared models, as well as amongst software, by computing distances between the vertices of the final meshes. Comparisons were also performed with models produced with the F6 SR scanner. These tasks were performed in the open source software CloudCompare with the Cloud-to-Cloud Hausdorff distances tool.

In the case study of the marble statue (Fig. 2), both spectral scenarios resulted in full digital reconstructions of the scene. Near-infrared image-based modeling produced sparser initial results than the visible spectrum modeling, however, dense results were of the same volumes. The NIR mesh contained a very low level of surface noise, especially in areas with less overlapping images, although the surface details had been equally preserved. In terms of geometric differences, the two models produced with Metashape Pro fluctuated by 0.7 mm $\pm$ 0.6 mm (Fig. 3), meaning a differentiation of less than four-times the GSD (1 RMS) or less than 2.5% of the smallest dimension of the object. In both cases, the distances between photogrammetric

models and the mesh produced from the F6 SR scanner were in the range of 0.8 $\pm$ 0.7 mm. Thus, considering the evaluated precision of the method, the models can be considered metrically identical and could be interchangeable for geometric recording and visualization on 1:5–1:10 scales-very common for detailed applications of decay mapping. Texturing results were of the same quality. As visible, shadows could not be effectively eliminated in either scenario.
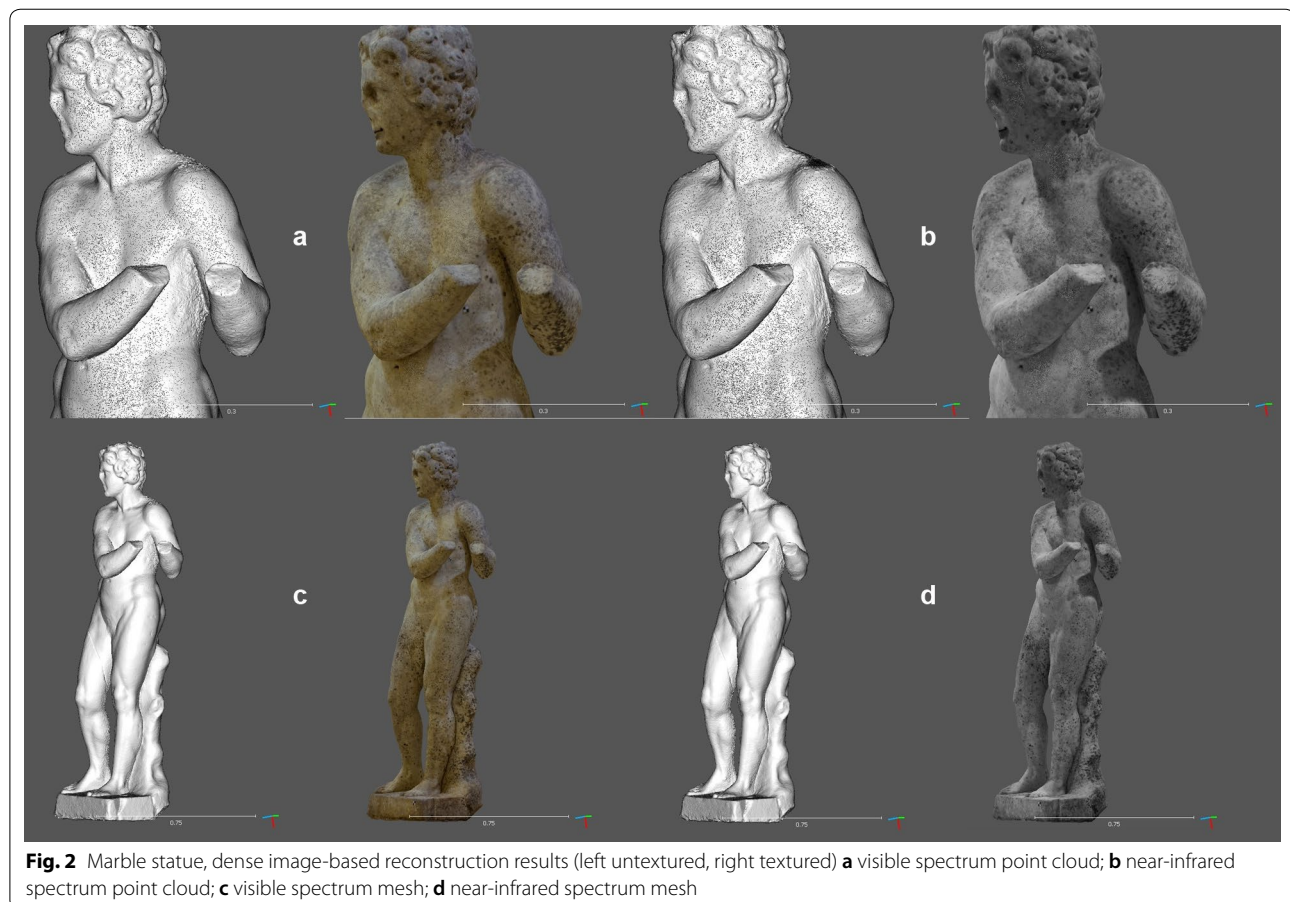
Observations on the 3D textured results for the marble statue (Fig. 4) showed that generally, higher near-infrared intensities correspond to healthier areas of the marble's surface, whereas lower intensities correspond to more deteriorated areas [32]. Thus, the NIR three-dimensional textured mesh could be used to roughly visualize in 3D the levels of decay on the statue. Undeniably, complete elimination of shadows would help optimize these results. We should further underline, that under no circumstance could this visualization substitute the detailed chemical characterizations. Although, for small, homogeneously shadowed areas of the same material, it can provide a fast approximation of the state-of-preservation and quantification of decay.

Regarding the case study of the lacquered screen panel, the mesh produced from NIR imagery by Metashape Pro was the only complete one and had a higher quality of surface detail. Metashape Pro produced about three-times denser results than Zephyr Aerial, but the imaging spectrum did not seem to have a significant effect on the density. However, the geometry of the 3D VIS models was problematic, highlighting the consequences of the glaring effect during visible spectrum data capturing (Fig. 5). Zephyr Aerial models were noisier than those produced by Metashape Pro. Distances between VIS and NIR geometries recorded with Metashape Pro were 1.0 mm $\pm$ 0.8 mm, meaning a differentiation of less than three-times the GSD (1 RMS), again, metrically very similar, considering the precision of the method. Still, distances computed for Zephyr Aerial were 2.2 mm $\pm$ 1.6 mm, which should be considered significant. Additionally, differences between the VIS models produced with different software ranged between 2 and 2.5 mm, the same as for the NIR models. As observed, the differences calculated between the 'perforated' Metashape Pro VIS model and the respective NIR model corresponded roughly to deficiencies and noise of the digitization in the visible spectrum (Fig. 6).

High-resolution modeling and texturing with near-infrared imagery did not only produce better topological results for the case study of the Chinese lacquerware but also enabled the identification of previous conservation interventions that are not recognizable on the

Table 2 Image-based reconstruction processing results

Case study	1		2				3				4			
Spectrum	VIS	NIR	VIS	NIR	VIS	NIR	VIS	NIR	VIS	NIR	VIS	NIR	VIS	NIR
Software	AMP	AMP	AMP	AMP	FZA	FZA	AMP	AMP	FZA	FZA	AMP	AMP	FZA	FZA
Sparse cloud														
Images aligned	180	180	22	22	22	22	40	40	40	40	60	60	60	60
Matching time [mm:ss]	05:16	14:46	00:34	00:29	04:02	04:02	01:16	01:07	07:08	07:14	01:55	01:23	43:56	11:59
Alignment time [mm:ss]	05:33	07:33	00:09	00:15	01:08	01:30	00:15	00:17	01:19	00:58	01:19	00:47	04:31	03:05
Tie points [1000]	454	151	94	87	28	28	81	71	43	44	420	236	132	123
Projections [1000]	1434	467	364	353	341	253	212	208	217	215	1270	643	803	667
Reprojection error [pixels]	0.65	0.72	0.58	0.69	0.82	0.83	0.90	0.93	1.08	1.10	0.35	0.48	0.47	0.50
Dense cloud														
Processing time [mm:ss]	1:43:36	1:57:41	08:17	08:52	16:53	18:00	07:07	07:48	37:19	29:27	14:27	15:15	42:16	43:49
Point count [1000]	14676	15122	5879	5394	1748	1840	25554	26542	6527	4276	10005	10082	3959	3960
Mesh														
Processing time [mm:ss]	08:35	15:14	01:41	01:22	03:35	03:35	13:40	14:33	08:49	09:02	02:58	02:58	03:34	03:21
Texture														
Processing time [mm:ss]	44:49	38:00	72:45	03:15	02:26	02:06	03:04	02:58	07:07	06:57	29:46	22:23	09:50	08:09
Overall														
Total time [h:mm:ss]	2:47:49	3:13:14	1:23:26	0:14:13	0:28:04	0:29:13	0:25:22	0:26:43	1:01:42	0:53:38	0:50:25	0:42:46	1:44:07	1:10:23

Fig. 2 Marble statue, dense image-based reconstruction results (left untextured, right textured) **a** visible spectrum point cloud; **b** near-infrared spectrum point cloud; **c** visible spectrum mesh; **d** near-infrared spectrum mesh

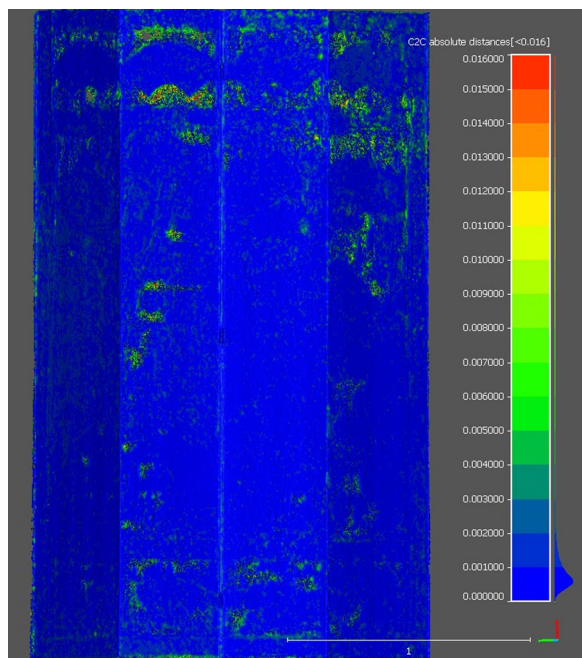Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 7 of 12
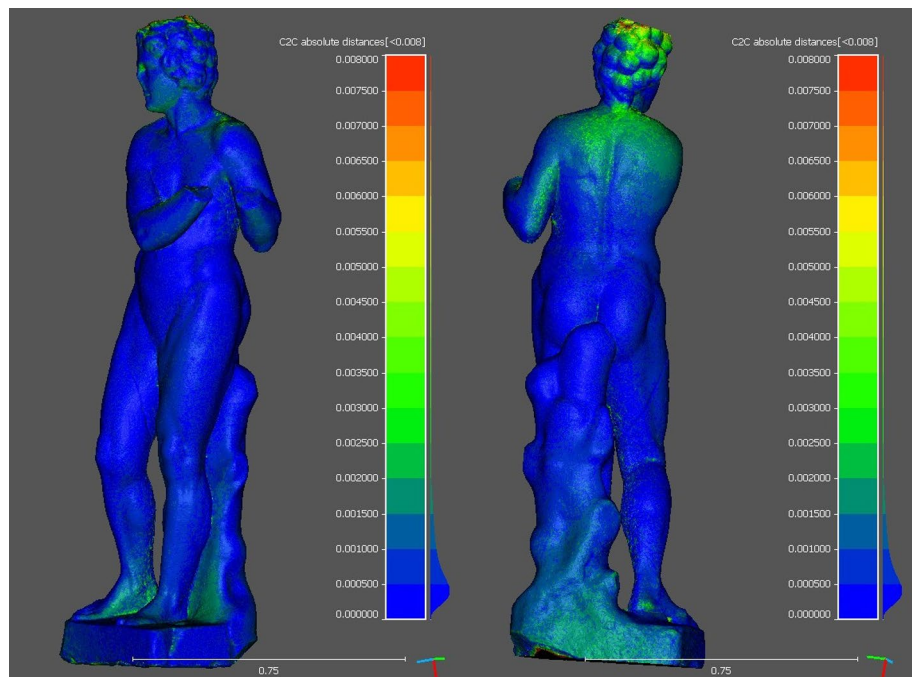
Fig. 3 Marble statue, geometric differences between visible and near-infrared spectra models (Hausdorff distances shown on a scale up to 8 mm, with 0.5 mm scalar field intervals)

Fig. 4 Marble statue, close-up comparison between visible (left), near-infrared (center) textures and 8-group classification mapping of near-infrared intensities (right)

visible textures and cannot be easily observed otherwise (Fig. 7).

In both cases, the distances between photogrammetric models and the meshes produced from the F6 SR scanner were in the range of 1.0 ± 1.0 mm.

For the case study of the wooden furniture part all, scenarios produced high-resolution results, with reconstruction errors of the same grade of magnitude (0.35–0.50 pixels). Although, final point clouds for Zephyr Aerial were approximately five times sparser.

Distances between VIS and NIR geometries recorded with Metashape Pro were 2.2 mm $\pm$ 1.4 mm, and 2.4 mm $\pm$ 2.0 mm with Zephyr Aerial. These differences can be interpreted as significant inaccuracies regarding the size of the pixel on the object (0.02 mm). However, if we consider the 2 ± 2 mm differences between the photogrammetric models and the models produced with F6 SR as the precision of the methodology, then they can be considered acceptable (Fig. 8). A percentage of this problem though can be attributed to the flat

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 8 of 12
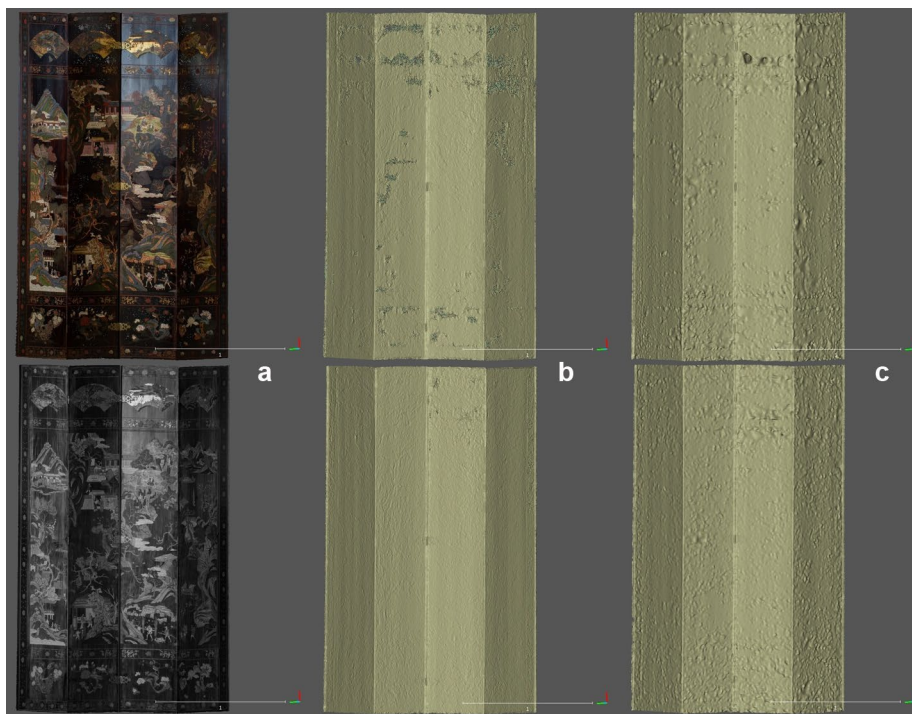
Fig. 5 Screen panel, dense image-based reconstruction results (top visible spectrum, bottom near-infrared spectrum), **a** AMP textured mesh; **b** AMP untextured mesh; **c** FZA untextured mesh

Fig. 6 Screen panel, geometric differences between visible and near-infrared spectra AMP models (Hausdorff distances shown on a scale up to 16 mm, with 1 mm scalar field intervals)

almost two-dimensional geometry of the object that does not favor the reconstruction. In addition to that, both 3D models from Zephyr Aerial included medium levels of noise. On the contrary, the respective 3D models from Metashape Pro had a very smooth surface while preserving most of the detail (Fig. 9).

The detail on acquired texturing results both for visible and near-infrared, was adequate for conservation-oriented observations. With the near-infrared textures we were able to successfully locate and measure cracks, paint defects and retouchings (Fig. 10).

Regarding the last case study, that of the stone sculpture of Christ crucified, all processing scenarios resulted in great detail and high texture-quality models (Fig. 11). Only the triangular 3D mesh produced with the NIR imagery from Zephyr Aerial contained a low magnitude of noise. Unlike to the other case studies, Zephyr Aerial produced sparser results than Metashape Pro both for the NIR and the VIS images. The initial reconstructions of the object were sparser in near-infrared but, dense reconstruction results were of very similar volume for VIS and NIR imagery in both software. The variation between VIS and NIR vertices ranged below 0.15 mm (0.08 mm $\pm$ 0.07 mm) for both software, a very accurate result considering the GSD of original photos, the precision of the method, the type
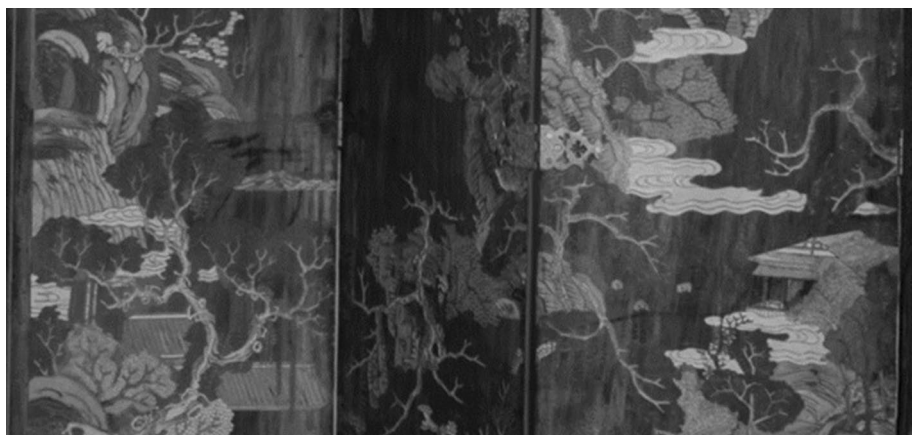
Fig. 7 Screen panel (detail), near-infrared model where previous interventions can be observed

of sensor, and the distances between the visible models, which had a 0.11 mm ± 0.14 mm (Fig. 12). The variation between both VIS and NIR photogrammetric models and the scanning model ranged below 0.27 mm. The very-high-resolution NIR textures made feasible the identification of deterioration and were further

Fig. 8 Wooden furniture part, geometric differences between visible and near-infrared spectra AMP models (top), geometric differences between visible spectrum models from two software (bottom; Hausdorff distances shown on a scale up to 8 mm, with 0.5 mm scalar field intervals)

exploited to roughly map the different levels of state of preservation (Fig. 13). This mapping could assist the precise work required by conservators on demanding restoration applications.

Conclusions

In this work, we constructed detailed models of historical artwork with high-resolution near-infrared textures, combining multi-view image 3D recording and near-infrared images captured with modified DSLR cameras, thus proving the feasibility of this integrated technique. Apart from the obvious advantage of reduced cost, compared to previously applied near-infrared three-dimensional acquisition methods, we also observed an increased versatility and convenience of implementation. We used two commercial software to compare the behavior of different Structure-from-Motion, Multiple-View-Stereo, and Meshing algorithmic implementations in near-infrared. Metashape Pro generally provided complete meshes, and less noisy than Zephyr Aerial, with higher density and slightly smaller statistical reconstruction errors. Furthermore, we constructed as a reference, detailed models of the same heritage case studies from visible imagery, acquired under the same conditions, and processed with similar parameters, to ensure comparability. We observed that the initial reconstruction results from all case studies were sparser in near-infrared, with slightly heightened reconstruction errors, but the dense results were of very similar volume and produced on comparable times. For the case study of a Chinese screen panel, which was from highly reflective material, digitization from near-infrared imagery even improved the reconstruction results, compensating for the glaring effects under visible light. Metric checks for all case studies proved that the combined 3D spectral technique
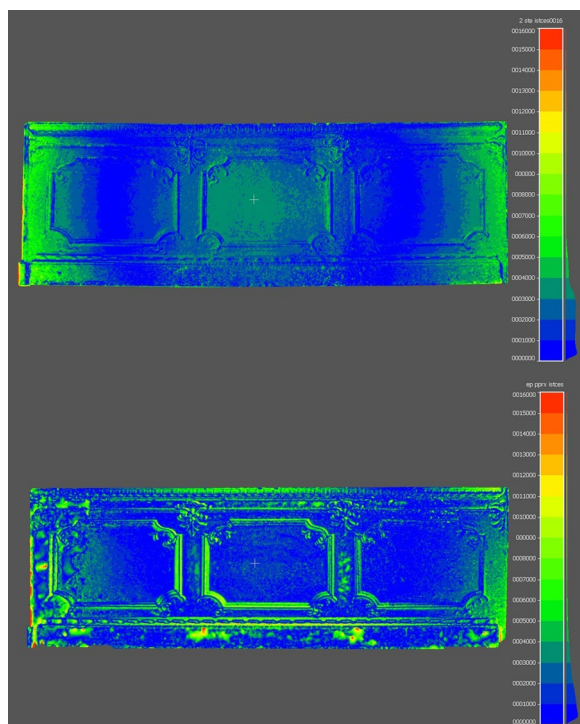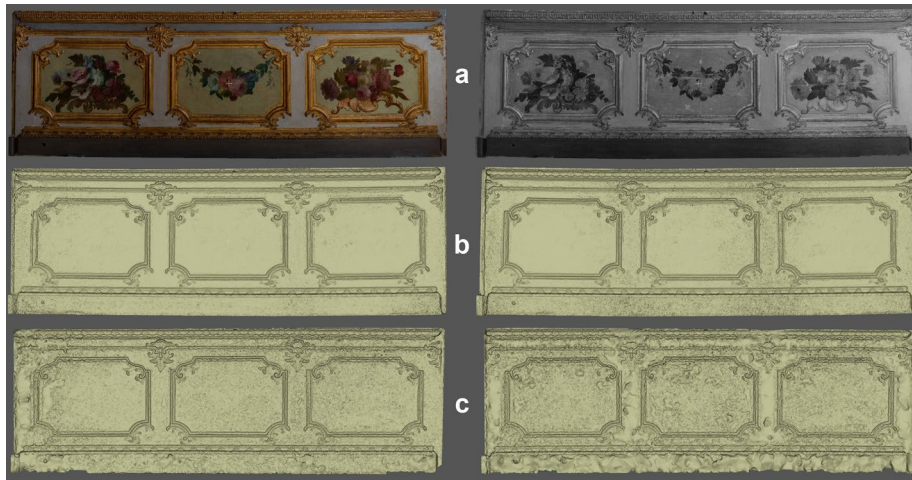
Fig. 9 Wooden furniture part, dense image-based reconstruction results (left visible spectrum, right near-infrared spectrum), **a** AMP textured mesh; **b** AMP untextured mesh; **c** FZA untextured mesh

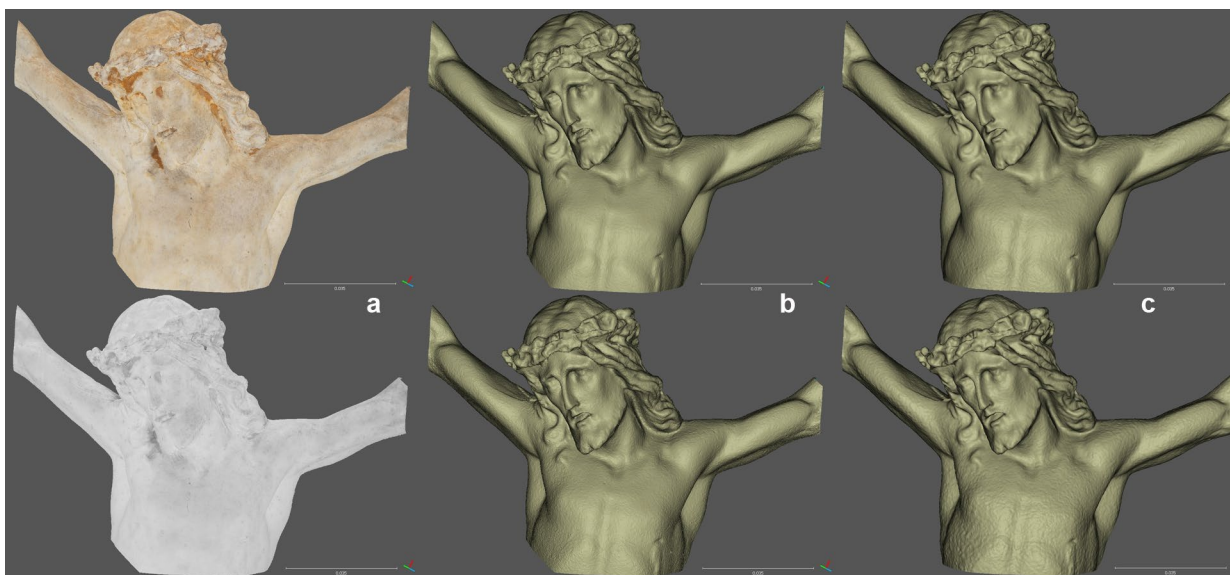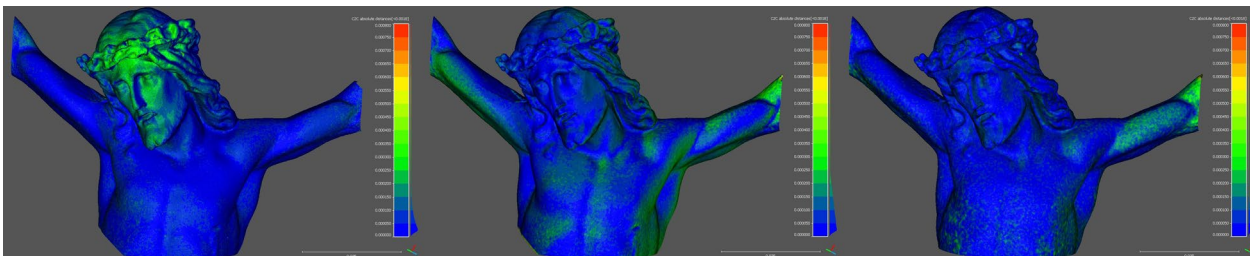**Fig. 10** Furniture part (detail), near-infrared texture orthophoto where cracks and re-touchments can be identified

Fig. 11 Sculpture of Christ, dense image-based reconstruction results (top visible spectrum, bottom near-infrared spectrum), **a** AMP textured mesh; **b** AMP untextured mesh; **c** FZA untextured mesh

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 11 of 12

Fig. 12 Sculpture of Christ, geometric differences between models, visible to near-infrared spectra for AMP (left); visible to near-infrared spectra for FZA (center); visible AMP to visible FZA (right); Hausdorff distances shown on a scale of max 0.8 mm, with 0.5 mm scalar field intervals)
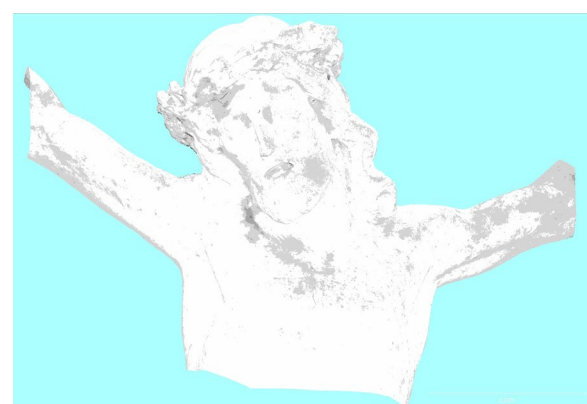
Fig. 13 Sculpture of Christ, 6-group classification mapping of near-infrared intensities from NIR AMP model

is also very accurate comparing to visible 3D image-based modeling, by taking into consideration sensors, lenses, capturing distances, and the precision of SfM/MVS based approaches. The high-resolution infrared textures improved the visibility of deterioration on the sculptures, potentially providing a valuable low-cost tool for three-dimensional decay mapping if certain light conditions are met during the data acquisition. In addition, they assisted the identification of defects and previous restoration works on the lacquerware and painted surfaces case studies, thus ascertaining the significance of the discussed approach towards heritage restoration and protection.

Abbreviations
3D: Three-dimensional; AMP: Agisoft Metashape Professional; CCD: Charge-coupled device; CMOS: Complementary metal-oxide-semiconductors; DSLR: Digital single-lens reflex camera; FZA: 3DFlow zephyr aerial; GSD: Ground sampling distance; MVS: Multiple-view-stereo; NIR: Near-infrared; RMS: Root square mean; SfM: Structure-from-Motion; VIS: VISible.

Authors' contributions
EA: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing—Original Draft, Visualization. AB: Resources, Writing—Review & Editing, Supervision. FR: Resources, Writing—Review & Editing, Supervision, Project administration, Funding acquisition. All authors read and approved the final manuscript.

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author details
[1] Department of Computer Science, Università degli Studi di Torino, Corso Svizzera 185, 10149 Turin, TO, Italy. [2] Fondazione Centro Conservazione e Restauro dei Beni Culturali 'La Venaria Reale', Via XX Settembre 18, 10078 Venaria Reale, TO, Italy. [3] Department of Architecture and Design, Politecnico di Torino, Viale Mattioli 39, 10125 Turin, TO, Italy.

References
1. Koutsoudis A, Vidmar B, Ioannakis G, Arnaoutoglou F, Pavlidis G, Chamzas C. Multi-image 3D reconstruction data evaluation. J Cult Herit. 2014;15(1):73–9. https://doi.org/10.1016/j.culher.2012.12.003.
2. Georgopoulos A, Stathopoulou EK. Data Acquisition for 3D Geometric Recording: State of the Art and Recent Innovations. In: Vincent ML,

Adamopoulos *et al. Herit Sci*   (2020) 8:53

Page 12 of 12

López-Menchero Bendicho VM, Ioannides M, Levy TE, editors. Heritage and archaeology in the digital age. Cham: Springer International Publishing; 2017. p. 1–26. https://doi.org/10.1007/978-3-319-65370-9_1.

3. McCarthy J. Multi-image photogrammetry as a practical tool for cultural heritage survey and community engagement. Archaeol Sci. 2014;43:175–85. https://doi.org/10.1016/j.jas.2014.01.010.

4. Cabrelles M, Blanco-Pons S, Carrión-Ruiz B, Lerma JL. From Multispectral 3D Recording and Documentation to Development of Mobile Apps for Dissemination of Cultural Heritage. In: Levy TE, Jones IWN, editors. Cyber-Archaeology and Grand Narratives. Cham: Springer International Publishing; 2018. p. 67–90. https://doi.org/10.1007/978-3-319-65693-9_5.

5. Georgopoulos A. Contemporary Digital Technologies at the Service of Cultural Heritage. In: Chanda B, Chaudhuri S, Chaudhury S, editors. Heritage Preservation. Singapore: Springer Singapore; 2018. p. 1–20. https://doi.org/10.1007/978-981-10-7221-5_1.

6. Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. 'Structure-from-Motion' photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology. 2012;179:300–14. https://doi.org/10.1016/j.geomorph.2012.08.021.

7. Fonstad MA, Dietrich JT, Courville BC, Jensen JL, Carbonneau PE. Topographic structure from motion: a new development in photogrammetric measurement. Earth Surf Process Landforms. 2013;38(4):421–30. https://doi.org/10.1002/esp.3366.

8. England Historic. Photogrammetric Applications for Cultural Heritage. Guidance for Good Practice. Swindon: Historic England; 2017.

9. Aicardi I, Chiabrando F, Lingua AM, Noardo F. Recent trends in cultural heritage 3D survey: the photogrammetric computer vision approach. J Cult Herit. 2018;32:257–66. https://doi.org/10.1016/j.culher.2017.11.006.

10. Kuchitsu N, Morii M, Sakai S, Unten H. Simple evaluation of the degradation state of cultural heritage based on multi-view stereo. Prog Earth Planet Sci. 2019;6:12. https://doi.org/10.1186/s40645-019-0260-7.

11. Robson S, Bucklow S, Woodhouse N, Papadaki H. Periodic photogrammetric monitoring and surface reconstruction of a historical wood panel painting for restoration purposes. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2004;XXXV:395–400.

12. Angheluță LM, Rădvan R. Macro Photogrammetry for the Damage Assessment of Artwork Painted Surfaces. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2019;XLII-2-W15:101–7. https://doi.org/10.5194/isprs-archives-xlii-2-w15-101-2019.

13. Fischer C, Kakoulli I. Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. Stud Conserv. 2006;51(Suppl 1):3–16. https://doi.org/10.1179/sic.2006.51.Supplement-1.3.

14. Liang H. Advances in multispectral and hyperspectral imaging for archaeology and art conservation. Appl Phys A. 2012;106(2):309–23. https://doi.org/10.1007/s00339-011-6689-1.

15. Delaney JK, Zeibel JG, Thoury M, et al. Visible and infrared imaging spectroscopy of Picasso's Harlequin Musician: mapping and identification of artist materials in situ. Appl Spectrosc. 2010;64(6):584–94. https://doi.org/10.1366/000370210791414443.

16. Bendada A, Sfarra S, Ibarra-Castanedo C, et al. Subsurface imaging for panel paintings inspection: a comparative study of the ultraviolet, the visible, the infrared and the terahertz spectra. Opto-Electron Rev. 2015;23(1):88–99. https://doi.org/10.1515/oere-2015-0013.

17. Miguel C, Bottura S, Ferreira T, Conde AF, Barrocas-Dias C, Candeias A. Unveiling the underprintings of a late-fifteenth-early-sixteenth century illuminated French incunabulum by infrared reflectography. J Cult Herit. 2019;40:34–42. https://doi.org/10.1016/j.culher.2019.05.014.

18. Cavaleri T, Buscaglia P, Lo Giudice A, et al. Multi and Hyperspectral Imaging and 3D Techniques for Understanding Egyptian Coffins. In: Strudwick H, Dawson J, editors. Ancient Egyptian Coffins: Past–Present–Future. Barnsley: Oxbow Books; 2019. p. 43–51. https://doi.org/10.2307/j.ctvh9w0cw.

19. Pronti L, Romani M, Verona-Rinati G, et al. Post-Processing of VIS, NIR, and SWIR multispectral images of paintings. New discovery on The Drunkenness of Noah, Painted by Andrea Sacchi, stored at Palazzo Chigi (Ariccia, Rome). Heritage. 2019;2(3):2275–86. https://doi.org/10.3390/heritage2030139.

20. Triolo PAM, Spingardi M, Costa GA, Locardi F. Practical application of visible-induced luminescence and use of parasitic IR reflectance as relative spatial reference in Egyptian artifacts. Archaeol Anthropol Sci. 2019;11(9):5001–8. https://doi.org/10.1007/s12520-019-00848-x.

21. Willard C, Gibson AP, Wade N. High-resolution visible and infrared imaging for large paintings a case study of Israel in Egypt by Poynter. In: Targowski P, Groves R, Liang H, editors. Optics for arts, architecture, and archaeology VII. Munich: SPI E; 2019. p. 110581F-1-110581F-12. https://doi.org/10.1117/12.2525714.

22. Zainuddin K, Majid Z, Ariff MFM, et al. 3D modeling for rock art documentation using lightweight multispectral camera. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2019;XLII-2/W9:787–93. https://doi.org/10.5194/isprs-archives-xlii-2-w9-787-2019.

23. Grifoni E, Legnaioli S, Nieri P, et al. Construction and comparison of 3D multi-source multi-band models for cultural heritage applications. J Cult Herit. 2018;34:261–7. https://doi.org/10.1016/j.culher.2018.04.014.

24. Webb EK, Robson S, MacDonald L, Garside D, Evans R. Spectral and 3D cultural heritage documentation using a modified camera. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2018;XLII-2:1183–90. https://doi.org/10.5194/isprs-archives-xlii-2-1183-2018.

25. Adamopoulos E, Rinaudo F, Bovero A. First assessments on heritage science oriented image-based modeling using low-cost modified and mobile cameras. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2019;XLII-2/W17:23–30. https://doi.org/10.5194/isprs-archives-xlii-2-w17-23-2019.

26. Mathys A, Jadinon R, Hallot P. Exploiting 3D multispectral texture for a better feature identification for cultural heritage. ISPRS Ann Photogramm Remote Sens Spatial Inf Sci. 2019;IV-2/W6:91–7. https://doi.org/10.5194/isprs-annals-iv-2-w6-91-2019.

27. Lanteri L, Agresti G, Pelosi C. A new practical approach for 3D documentation in ultraviolet fluorescence and infrared reflectography of polychromatic sculptures as fundamental step in restoration. Heritage. 2019;2:207–15. https://doi.org/10.3390/heritage2010015.

28. Pamart A, Guillon O, Faraci S, et al. Multispectral photogrammetric data acquisition and processing for wall paintings studies. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2017;XLII-2-W3:559–66. https://doi.org/10.5194/isprs-archives-xlii-2-w3-559-2017.

29. Toldo R, Fantini F, Giona L, et al. Accurate multiview stereo reconstruction with fast visibility integration and tight disparity bounding. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2013;XL-5/W1:243–9. https://doi.org/10.5194/isprsarchives-xl-5-w1-243-2013.

30. Alsadik B, Gerke M, Vosselman G, Daham A, Jasim L. Minimal camera networks for 3D image based modeling of cultural heritage objects. Sensors. 2014;14(4):5785–804. https://doi.org/10.3390/s140405785.

31. De Reu J, Plets G, Verhoeven G, et al. Towards a three-dimensional cost-effective registration of the archaeological heritage. J Archaeol Sci. 2013;40(2):1108–21. https://doi.org/10.1016/j.jas.2012.08.040.

32. Rahrig M, Drewello R, Lazzeri A. Opto-technical monitoring—a standardized methodology to assess the treatment of historical stone surfaces. Int Arch Photogramm Remote Sens Spatial Inf Sci. 2018;XLII-2:945–52. https://doi.org/10.5194/isprs-archives-xlii-2-945-2018.

Publisher's Note