Hyperspectral Analysis of Pigment Identification and Abundance Inversion in the Dome of China’s Yungang Grotto 7

Junlin Yan [1,2], Shuqiang Lyu [1,2], Bo Ning [3], Lihong Li [3], Miaole Hou [4*]
[1] School of Geomatics and Urban Spatial Informatics, Beijing University of Civil Engineering and Architecture, No.15 Yongyuan Road, Daxing District, Beijing, China -2108570424056@stu.bucea.edu.cn, lvshuqiang@bucea.edu.cn
[2] Beijing Key Laboratory for Architectural Heritage Fine Reconstruction & Health Monitoring, No.15 Yongyuan Road, Daxing District, Beijing, China
[3] Yungang Research Institute, Datong 037007, China-13233616482@163.com, ningbo007@163.com
[4] Chang’an University, No.126 Yanta Road, Yanta District, Xi’an 710064, Shaanxi, China, houmiaole@bucea.edu.cn

Keywords: Yungang Grottoes, Hyperspectral, Spectral Unmixing, Abundance Inversion

Abstract:

Most of the grotto temples have undergone long-term weathering and multiple repainting campaigns, so accurate identification of the composition and spatial abundance of surface pigments is an important foundation for pigment characterization and conservation research. This study focuses on the dome of Yungang Grotto 7. Data were acquired using a three-dimensional (3D) hyperspectral multimodal digital acquisition system and the Analytical Spectral Devices (ASD) field spectroradiometer. The workflow consisted of two stages: pigment identification and abundance inversion. In the pigment identification stage, a normalized weighted identification method integrating Spectral Angle Mapper (SAM) and the Normalized Difference Spectral Index (NDSI) was proposed based on mineral pigment reflectance curves measured by the ASD field spectroradiometer. In the abundance inversion stage, Fully Constrained Least Squares (FCLS) was applied to estimate pigment proportions in mixed pixels under non-negativity and sum-to-one constraints. The results show that the green pigments are most likely malachite and Paris green, the red pigments are most likely hematite and laterite, and the black pigment is most likely carbon black. The interwoven distribution of Paris green and traditional mineral pigments provides material-science evidence for modern repainting and restoration in this area. Nonlinear mixing may occur on rough and weathered grotto surfaces. However, under the current data conditions, its influence on abundance inversion remains unclear. Therefore, Kernel Fully Constrained Least Squares (K-FCLS) was additionally introduced as a reference nonlinear model for qualitative comparison with FCLS.

1 Introduction

As an important component of China’s ancient cultural heritage, the polychrome paintings in grotto temples, especially those in Yungang Grottoes dating to the Northern Dynasties (between AD 460 and AD 524), possess exceptionally high historical and artistic value (Fan et al., 2025). However, under the complex effects of natural weathering, particulate deposition, and human repainting, the pigments on the surface of the grotto have suffered severe fading and loss of pigment layer (Fan et al., 2024). At present, the conservation of grotto temples urgently requires a transition from purely visual qualitative assessment to rigorous high-precision quantitative analysis.

Conventional pigment analysis methods, such as portable X-ray fluorescence (XRF) and Raman spectroscopy, are capable of high-precision compositional identification (Cosano et al., 2019). However, due to their inherent single-point testing nature, these methods cannot provide continuous areal information on the spatial distribution of pigments. Owing to its integrated spatial-spectral capability and high spectral resolution, hyperspectral imaging can acquire continuous narrow-band spectral data and has become a core non-destructive means for investigating pigment distribution over large cultural heritage surfaces (Cucci et al., 2020). In recent years, hyperspectral technology has been used for the qualitative identification and quantitative spectral unmixing of pigments in cultural heritage.

In quantitative pigment analysis based on hyperspectral technology, abundance inversion is a crucial step. In hyperspectral spectral unmixing applications, Grabowski et al. (2018) combined automatic endmember extraction with spectral library matching to achieve automatic pixel-level pigment identification. Zapiola et al. (2025) pointed out that Fully Constrained Least Squares (FCLS), based on the Linear Mixing Model (LMM), demonstrates strong suitability for quantitative inversion because it strictly satisfies the physical constraints of non-negativity and abundance sum-to-one. This provides a solid theoretical and methodological basis for the use of FCLS in this study. However, in severely weathered grotto environments, the intimate mixing of pigment particles and the extreme roughness of the surface often induce a certain degree of nonlinear mixing effects (Rohani et al., 2018). Although non-linear models are more accurate in simulating light scattering by microscopic particles, their extremely high computational complexity limits their scope of application; for the study and quantitative analysis of large-scale grottoes, models based on LMM (such as FCLS) remain the most widely used and practical option. In addition, many existing unmixing studies directly apply algorithms to blindly extract endmembers from images without solid material-based validation, which greatly compromises the reliability of inversion results in complex scenarios.

To address the issues, this study focuses on the dome of Grotto 7 at the Yungang Grottoes. This study follows an overall workflow of qualitative identification followed by quantitative inversion. Firstly, during the pigment identification stage, a normalized weighted identification method utilizing Spectral Angle Mapper (SAM) and Normalized Difference Spectral Index (NDSI) was employed to determine the specific types of pigments. Secondly, during the abundance inversion stage, the identification results were input into the FCLS model as known endmembers for abundance inversion on large areas of painted surfaces. In addition, Kernel Fully Constrained Least Squares (K-FCLS) was introduced as a reference nonlinear model for qualitative comparison with FCLS. The findings provide support for the identification of pigment types in weathered grotto paintings and offer a scientific basis for the preventive conservation and deterioration monitoring of contemporary grottoes in the surrounding areas.

2 Methodology

2.1 Study Area

The Yungang Grottoes were carved during the Northern Wei Dynasty in the 5th century AD and are an outstanding example of Northern Wei Dynasties grotto art. As a typical representative grotto of the middle phase of Yungang, Grotto 7 preserves relatively abundant remains of ancient polychrome paintings in its dome area. Unlike traditional earthen-layer murals preserved in indoor museum environments, the polychrome paintings in this grotto are directly attached to a sandstone substrate in a semi-open environment (Fan et al., 2024). Under the combined effects of long-term natural weathering, particulate deposition, and repeated human repainting, the surface pigments have become complex and difficult to be identified. In this study, the central area of the southern dome zone of Yungang Grotto 7 was selected as the study area, which covers approximately 2.3 m², and a total of 18 hyperspectral images were acquired. Due to limitations in on-site acquisition conditions, two scenes from the study area were not included in the final mosaic, resulting in missing information in some local areas. Nevertheless, the mosaic result still preserved the main pigment distribution of the study area well, providing reliable support for subsequent pigment identification and abundance distribution analysis.

2.2 Instruments and Equipment

(1) Hyperspectral image data were acquired using a 3D hyperspectral multimodal digital acquisition system. The system integrates a hyperspectral camera, a 3D sensor, and a constant-temperature broad-spectrum LED cold light source. The hyperspectral camera covers the spectral range of 400-1000 nm, with 300 bands, a spectral resolution better than 2.8 nm, and a spatial resolution of over 4 million pixels. The average point spacing of the 3D sensor point cloud is approximately 0.6 mm.

(2) Point spectral data were collected using the FieldSpec 4 portable field spectroradiometer manufactured by ASD (Analytical Spectral Devices, USA), with a spectral range of 350-2500 nm and a total of 2151 bands. The spectral resolution is 3 nm in the 350-1000 nm range and 8 nm in the 1001-2500 nm range.

(3) Micro-Raman spectroscopy was performed using an ALMEGA micro-confocal laser Raman spectrometer manufactured by Thermo Nicolet (USA) to analyze the composition of pigment samples. The laser power was 50 mW at 780 nm and 25 mW at 532 nm.

(4) Scanning electron microscopy (SEM) was conducted using a TESCAN MIRA LMS instrument from the Czech Republic, together with a Smartedx energy-dispersive spectrometer (EDS) for semi-quantitative analysis of the elements contained in the pigments. During testing, small pigment samples were mounted on conductive adhesive, and the sample surfaces were platinum coated.

2.3 Sample Collection

The spectral acquisition of the mural is shown in Figure 1, and the information on pigment samples analyzed by micro-Raman spectroscopy and SEM-EDS is presented in Table 1.

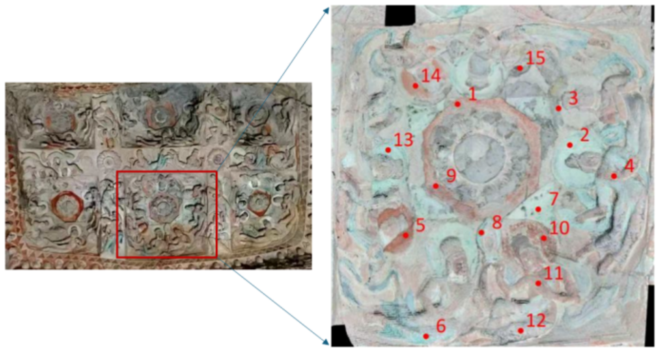


Figure 1. Spectral sampling locations in the dome of Yungang Grotto 7

No.	Sample ID	Sample Color	Measurement Points	Sampling Location
1	YGC-QD-1	Green	8,13	Ribbon above the leg in the central southern area
2	Not Sampled	Green	2,3,6,7	
3	YGC-QD-2	Red (light)	4,5	Red triangular border in the southeastern area
4	YGC-QD-3	Red (dark)	1,9,10,11,14	Buddha foot at the lower right corner of the central opening on the southern side
5	YGC-QD-4	Black	12,15	Lower right side of the Buddha head in the southeastern area

Table 1. Pigment samples from Yungang Grotto 7

2.4 Methods

2.4.1 Pigment Identification: This study employs a spectral matching method to compare the acquired spectra of target pixels with the pigment spectral library established in the laboratory. Considering that the undulating curved surface of the grotto dome easily leads to uneven illumination and thus affects the stability of spectral reflectance, we developed a weighted identification method integrating SAM and NDSI to improve the accuracy and stability of pigment identification. Based on the hyperspectral identification results, micro-Raman spectroscopy was further employed to obtain high spatial resolution molecular vibrational spectral information of the pigments, while SEM-EDS was used to analyze the elemental composition of the pigment surfaces, thereby finally determining the pigment types on the dome surface.

SAM represents the target spectrum and the reference spectrum as vectors in a multidimensional space and characterizes their spectral shape similarity by calculating the angle between them. The specific formula is as follows:

$$SAM(x,y) = arccos\left(\frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}\right) \quad (1)$$

where x_i is the spectrum of the target pixel, y_i is the endmember spectrum in the standard spectral library, and n is the total number of bands.

NDSI performs a normalized calculation on the reflectance of the target spectrum at two characteristic bands, thereby expressing its spectral features in a dimensionless form, with values generally ranging from [-1, 1]. The specific formula is as follows:

$$NDSI(x_i, x_j) = \frac{x_i - x_j}{x_i + x_j} \tag{2}$$

where x_i and x_j represent the reflectance values of the target spectrum at the i-th and j-th characteristic bands, respectively. The characteristic bands were selected mainly according to the differential features of the standard spectra of different pigments in key reflection peaks, absorption valleys, and slope-variation regions, and were further determined in combination with the pre-experimental identification results.

In this study, the above two indices were first normalized using min-max normalization, and then an integrated matching index was constructed as follows:

$$Index = 0.7 \times SAM_{norm} + 0.3 \times NDSI_{norm} \tag{3}$$

where the weights of SAM_{norm} and $NDSI_{norm}$ were set to 0.7 and 0.3, respectively. This weighting scheme emphasizes the role of spectral shape features in pigment identification, while retaining the auxiliary discriminative capability of reflectance-value differences. By selecting the endmember with the minimum integrated identification index as the optimal matching result, the pigment type of the target can be identified, thereby providing reliable endmember input for the subsequent FCLS-based abundance inversion.

2.4.2 Abundance Inversion: FCLS is one of the classical linear unmixing methods in hyperspectral spectral unmixing and can rapidly obtain the abundance information of each endmember. This method is based on the LMM and assumes that the observed spectrum can be expressed as a linear combination of multiple endmember spectra, as follows:

$$\boldsymbol{x}_n = \sum_{p=1}^{P} a_{np} \boldsymbol{m}_p + \boldsymbol{\varepsilon}_n \tag{4}$$

$$a_{np} \geq 0, \sum_{p=1}^{P} a_{np} = 1 \tag{5}$$

where $\boldsymbol{x}_n$ is the reflectance of the n-th pixel in the hyperspectral image, P is the number of endmembers, $\boldsymbol{m}_p$ is the endmember reflectance of the p-th endmember, a_{np} is the abundance of the p-th endmember in the n-th pixel, and $\boldsymbol{\varepsilon}_n$ is the noise term. By applying the FCLS algorithm on a pixel-by-pixel basis to solve for the optimal abundance matrix, abundance maps of the spatial distribution of pigments on the grotto surface can ultimately be generated.

K-FCLS is a nonlinear unmixing method extended from FCLS. By introducing a kernel function, the original spectra are mapped into a high-dimensional feature space, which improves the representation of nonlinear mixing relationships. The model can be expressed as:

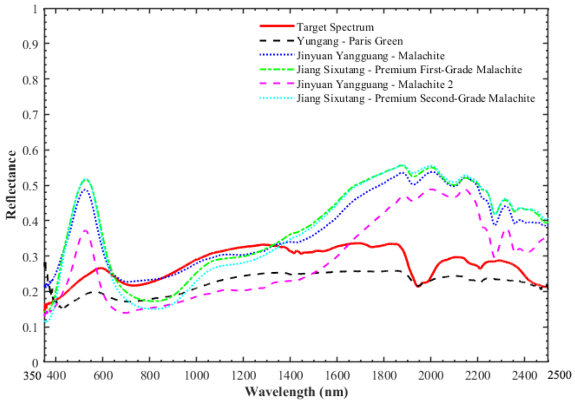
$$\Phi(x_n) = \sum_{p=1}^{P} a_{np} \, \Phi(m_p) + \varepsilon_n \tag{6}$$

where $\Phi()$ denotes the nonlinear mapping induced by the kernel function, x_n is the spectral vector of the n-th pixel in the hyperspectral image, P is the number of endmembers, m_p is the spectral reflectance of the p-th endmember, a_{np} denotes the abundance of the p-th endmember in the n-th pixel, and ε_n is the error term. By solving the K-FCLS model pixel by pixel, the abundance distribution of each endmember in the study area can be obtained, thereby generating pigment abundance maps.

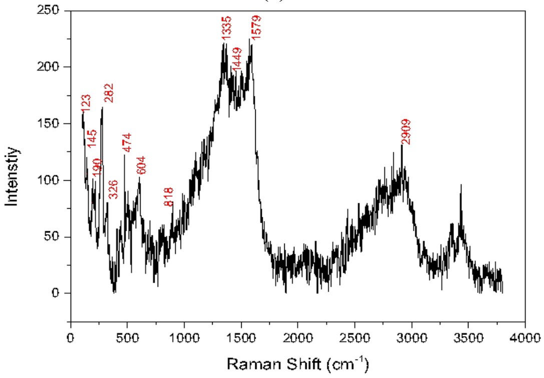
3 Results and Discussion

3.1 Pigment identification

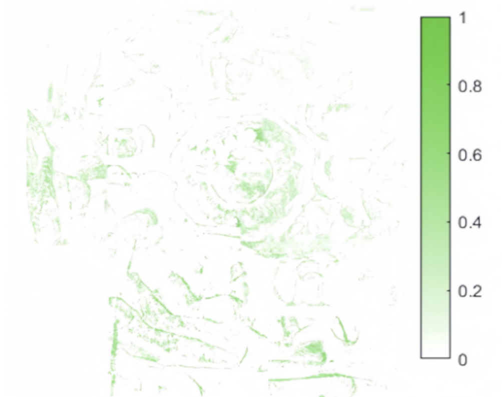
First, SAM and NDSI were used to identify the pigment types at the measured points, and the results were cross-validated with Raman spectroscopy and SEM-EDS. Then, FCLS was employed to invert the abundance distribution of each pigment in the hyperspectral images. Due to the limitations of on-site sampling conditions, only one sample was collected from a green painted area in this study. However, during hyperspectral data acquisition, multiple green points within the study area were sampled relatively uniformly. Subsequent spectral processing and analysis showed that the reflectance spectra of these green points could be divided into two categories. Among them, the green pigment YGC-QD-1 was distributed at points 8 and 13, while the other type of green pigment was distributed at points 2, 3, 6, and 7. YGC-QD-1 is light green and is mainly distributed in the halo of the figure and along the inner-ring boundary of the central circular motif, whereas the other green pigment constitutes the main visual component of the green painted areas on the dome and is widely distributed along the outer contour of the central motif, the interwoven ornamental bands, and the ribbons of the figures below. As shown in Figure 2(a), based on the spectral identification results, YGC-QD-1 was identified as Paris green. The Raman spectral results are given in Figure 2(b), obvious characteristic peaks can be observed at 123, 145, 190, 282, 326, 474, 604, 818, 1335, 1449, 1579, and 2909 cm^{-1}, based on which it can be further confirmed as Paris green. Moreover, baseline drift is present at some test points, indicating that the sample may be associated with organic components. As a modern synthetic pigment, the co-occurrence of Paris green and organic binding materials indicates that the dome of Yungang Grotto 7 may have undergone repainting after the 19th century. The corresponding abundance inversion result for Paris green is shown in Figure 2(c). According to the hyperspectral identification result shown in Figure 2(d), the pigment in this area was identified as malachite, and the corresponding abundance inversion result is shown in Figure 2(e). Malachite and Paris green appear simultaneously on the dome surface. The interwoven spatial distribution makes spectral mixing at the pixel scale more complex and also increases the difficulty of abundance inversion and fine discrimination.



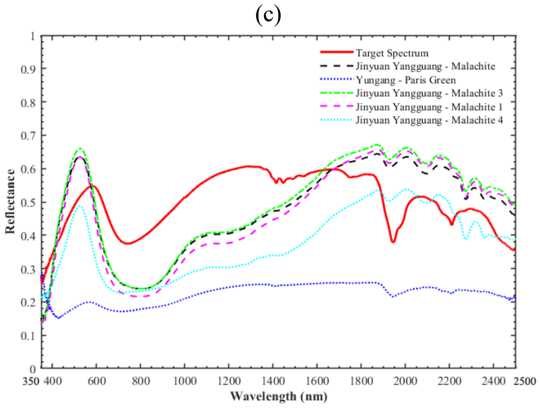
(a)



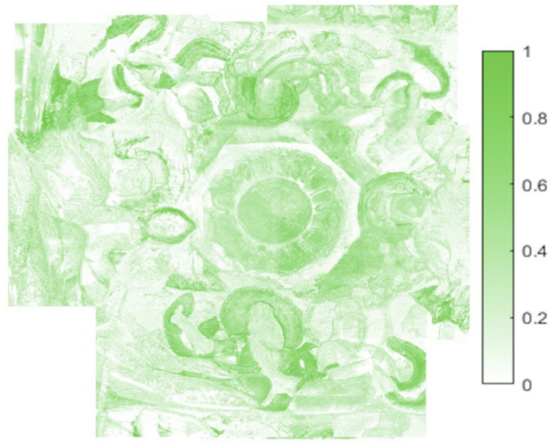
(b)



(c)



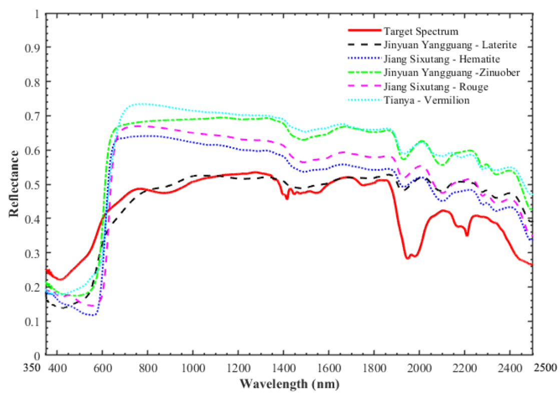
(d)



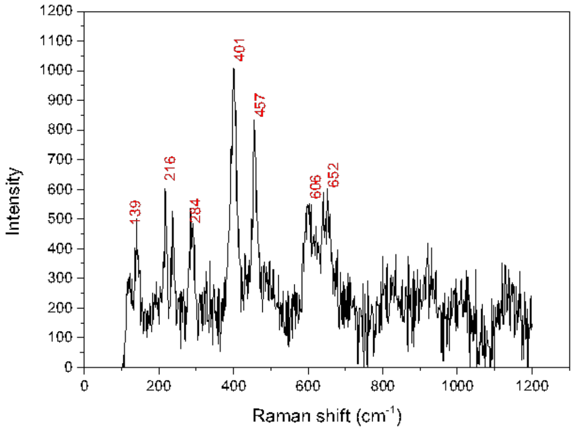
(e)

Figure 2. Analysis results of green pigments

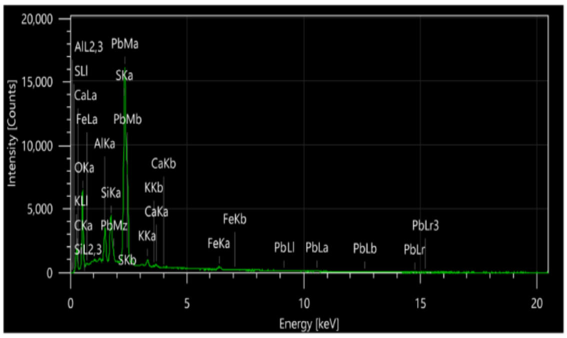
Two reddish pigments are present on the dome surface. The light red pigment YGC-QD-2, corresponding to points 4 and 5, and the dark-red pigment YGC-QD-3, corresponding to points 1, 9, 10, 11, and 14. The two kinds of pigment show spatial overlap. According to the hyperspectral identification result in Figure 3(a), YGC-QD-2 was preliminarily identified as hematite. Further combined with the Raman spectral analysis in Figure 3(b), obvious characteristic peaks can be observed at 139, 216, 284, 410, 457, 606, and 652 cm^{-1}, which are generally consistent with the Raman characteristic peaks of Fe_2O_3. In addition, baseline drift is present at some test points, indicating that the sample may be associated with organic components. The SEM-EDS result in Figure 3(c) shows that Fe and Pb were the main detected elements in this pigment, indicating that its principal components are Fe_2O_3 and lead oxides. The corresponding abundance inversion result is shown in Figure 3(d). Fe_2O_3 and hematite are closely related in cultural heritage materials science (Stuart and Thomas, 2017).



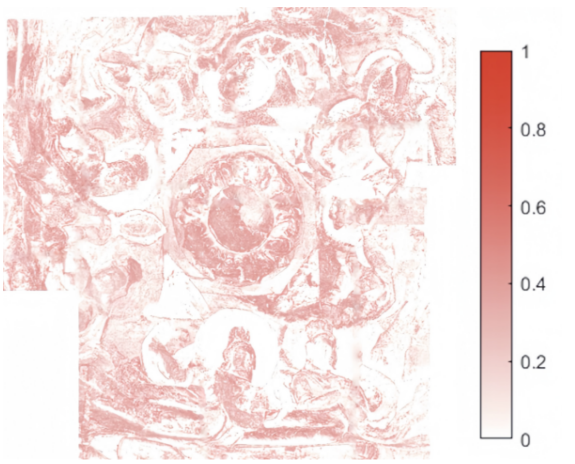
(a)



(b)



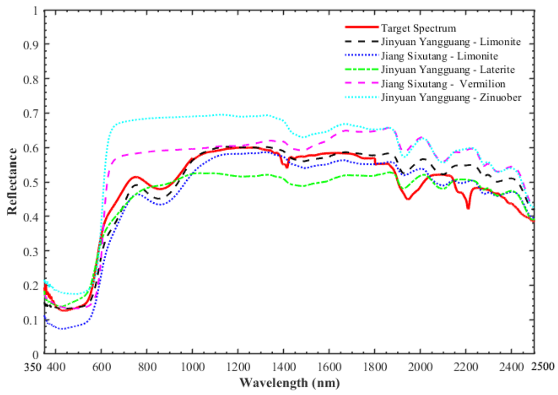
(c)



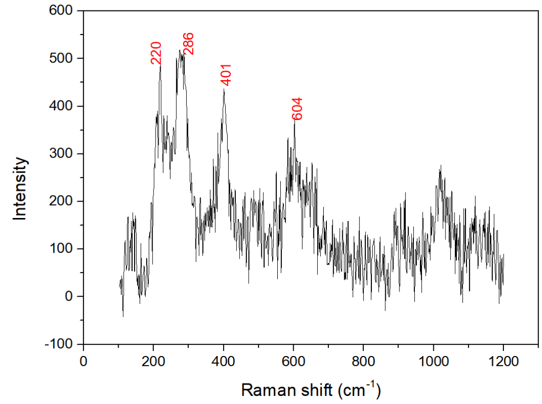
(d)

Figure 3. Analysis results of hematite pigment

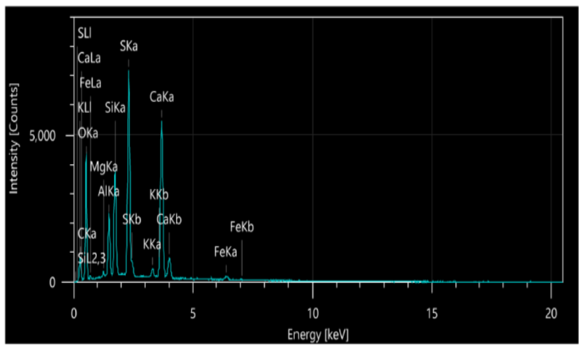
As shown in Figure 4(a), YGC-QD-3 was identified as laterite. As shown in Figure 4(b), the Raman spectroscopy revealed obvious peaks at 220, 286, 401, and 604 cm^{-1}, which are generally consistent with the Raman spectral characteristics of Fe_2O_3. The SEM-EDS result in Figure 4(c) did not reveal distinct chromophore pigment minerals. The main detected elements were Ca, S, Si, Al, and O, indicating the presence of clay minerals and gypsum. Laterite is a pigment made from iron-rich natural red earth or crushed red rock. It is a mixture that contains not only red chromophore components, but also a significant amount of impurities, such as clay, quartz, and mica. The corresponding abundance inversion result is shown in Figure 4(d).



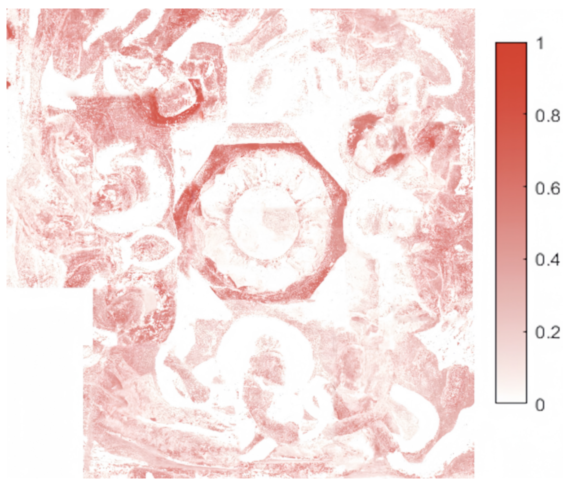
(a)



(b)



(c)

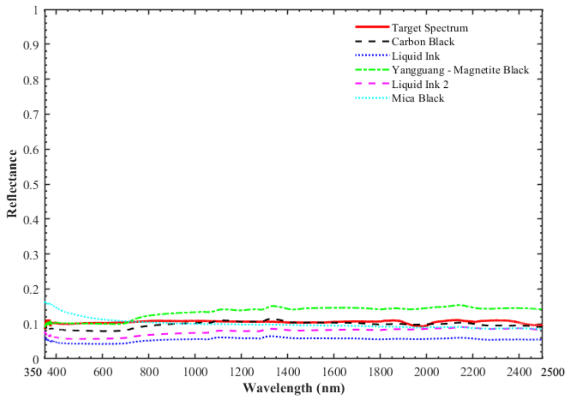


(d)

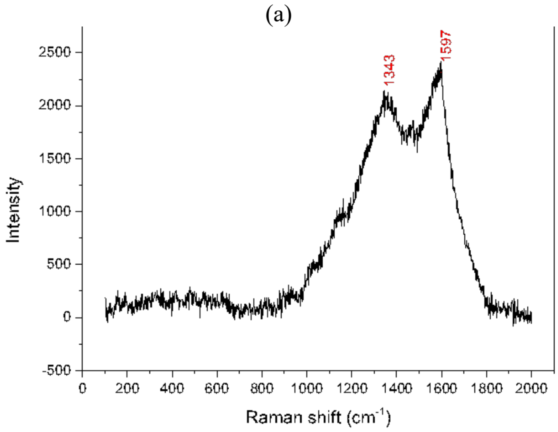
Figure 4. Analysis results of laterite pigment

In summary, the red pigments in this area are macroscopically identified as the pigment hematite and laterite, representing different color grades used in ancient times. However, at the microscopic level, Raman spectroscopic analysis indicates that they are all Fe_2O_3.

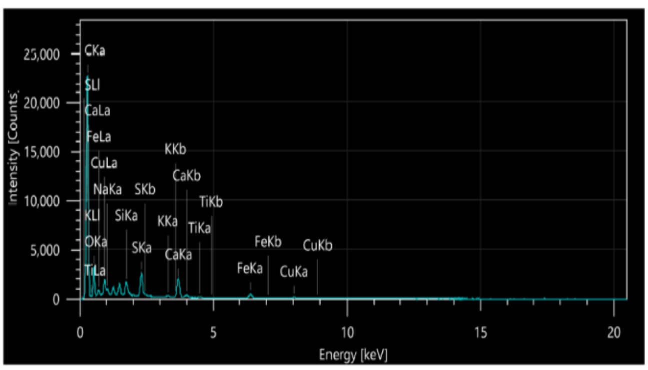
The black pigment YGC-QD-4, corresponding to points 12 and 15, mainly exhibits typical linear and local spot-like features. As shown in Figure 5(a), it was identified as carbon black. In Figure 5(b), Raman spectroscopy revealed obvious peaks at 1343 and 1597 cm^{-1}, which are generally consistent with the Raman spectral characteristics of carbon black. The SEM-EDS result in Figure 5(c) shows that the pigment is mainly composed of elemental C, indicating carbon black (C). The corresponding abundance inversion result is shown in Figure 5(d).



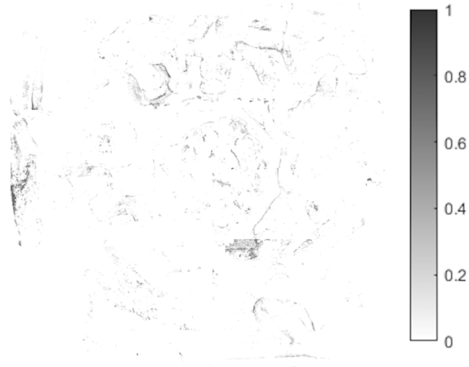
(a)



(b)



(c)



(d)

Figure 5. Analysis results of carbon black pigment

3.2 Analysis of Abundance Inversion Results

To examine whether the use of a nonlinear model would lead to substantial differences in the retrieved pigment distribution patterns under the current data conditions, a local area of the mural was selected for qualitative comparison using both K-FCLS and FCLS. Figure 6 shows a local true-color image of the mural, whilst Figures 7 to 11 present comparative diagrams of the abundance inversion for different pigments.

Figure 6. Local true-color image of the mural

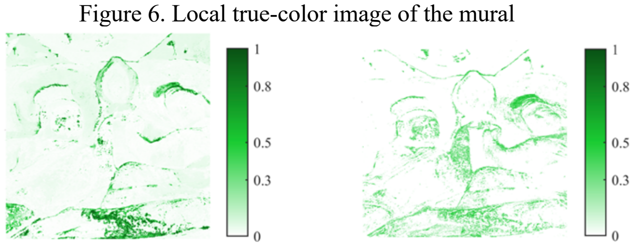


Figure 7. Abundance map of pigment YGC-QD-1

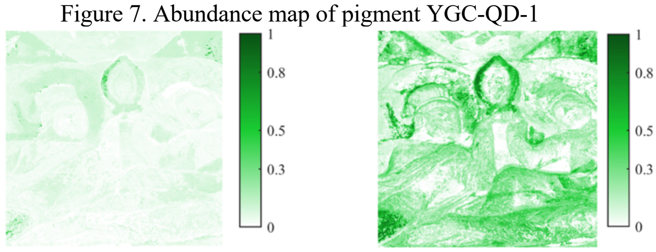


Figure 8. Abundance map of malachite pigment

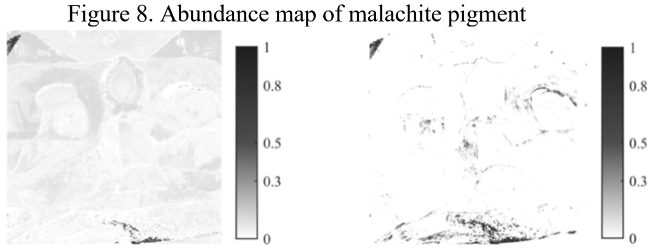


Figure 9. Abundance map of pigment YGC-QD-2

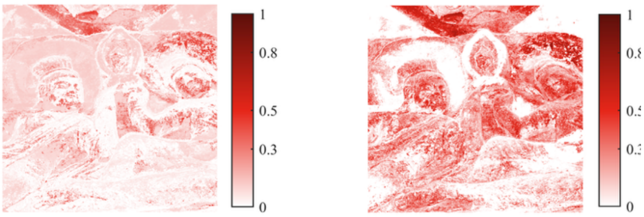


Figure 10. Abundance map of pigment YGC-QD-3

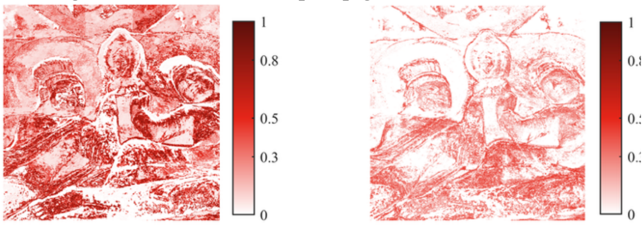


Figure 11. Abundance map of pigment YGC-QD-4

Comparison of the pigment abundance inversion maps indicates that, because the data were acquired from a three-dimensional sculptural surface, some black regions are attributable to shadows cast by the sculpture. The abundance inversion results for the red and green pigments are generally consistent in both models. In addition, FCLS required less computation time than K-FCLS for the same study area, indicating better efficiency for large-area hyperspectral abundance inversion. Therefore, under the current data conditions, FCLS was adopted in this study as a practical model for large-area pigment abundance inversion.

4 Conclusions

This study focused on the dome of Yungang Grotto 7. Data were acquired using a 3D hyperspectral multimodal digital acquisition system and an ASD portable spectroradiometer. Based on the acquired hyperspectral and point spectral data, a two-stage analytical framework of pigment identification and abundance inversion was established. In the pigment identification stage, a normalized weighted identification method integrating SAM and NDSI was employed. In the abundance inversion stage, the identified pigments were introduced into the FCLS model as known endmembers, thereby achieving spatial abundance inversion of pigments in the dome area. Under complex weathering conditions, the combined use of SAM and NDSI can effectively improve the reliability and stability of pigment identification in weathered grotto polychrome paintings. The results show that the green pigments in the study area are mainly malachite and Paris green, the red pigment is mainly Fe_2O_3, and the black pigment is mainly carbon black. In particular, the interwoven distribution of Paris green and traditional mineral pigments provides important evidence for later repainting and restoration in this area. A qualitative comparison with K-FCLS showed that the two models yielded generally similar visual distribution patterns for the main pigments, whereas FCLS had higher computational efficiency. Under the current data conditions, FCLS was therefore adopted for large-area hyperspectral abundance inversion in this study.

Overall, under the current data conditions, the linear model can better characterize the surface pigment distribution of the dome of Yungang Grotto 7. This study provides an effective technical pathway for quantitative pigment mapping of weathered grotto polychrome paintings and offers methodological support for the conservation research of grotto polychrome heritage. Owing to the limitations of on-site acquisition conditions, further validation and optimization are still needed in the future under more complex weathering conditions and across a wider range of pigment systems.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No.4271356 and No.42171444).

References

Cosano, D., Esquivel, D., Pérez, A., et al., 2019. Characterization of wall paintings from the caliphal baths of Cordoba (Spain) by X-ray diffraction and Raman microspectroscopy. Analytical Letters, 52(3), 411–422. https://doi.org/10.1080/00032719.2018.1470186

Cucci, C., Picollo, M., Chiarantini, L., Uda, G., Fiori, L., De Nigris, B., Osanna, M., 2020. Remote-sensing hyperspectral imaging for applications in archaeological areas: non-invasive investigations on wall paintings and on mural inscriptions in the Pompeii site. Microchemical Journal. https://doi.org/10.1016/j.microc.2020.105082

Fan, X., Cui, J., Wang, S., Tai, L., Guo, J., Yan, H., 2024. Analysis of pigments unearthed from the Yungang Grottoes archaeological excavations. Minerals, 14, 221. https://doi.org/10.3390/min14030221

Fan, X., Xie, H., Ma, R., et al., 2025. The ancient renovation activities of Yungang Grottoes: mutual verification of carbon-14 dating and archaeological investigation. Archaeological and Anthropological Sciences, 17, 32. https://doi.org/10.1007/s12520-025-02160-3

Grabowski, B., Masarczyk, W., Głomb, P., et al., 2018. Automatic pigment identification from hyperspectral data. Journal of Cultural Heritage, 31, 1–12. https://doi.org/10.1016/j.culher.2018.01.003

Rohani, N., Pouyet, E., Walton, M., et al., 2018. Nonlinear unmixing of hyperspectral datasets for the study of painted works of art. Angewandte Chemie, 57(34), 10910–10914. https://doi.org/10.1002/anie.201805135

Stuart, B.H., Thomas, P.S., 2017. Pigment characterisation in Australian rock art: a review of modern instrumental methods of analysis. Heritage Science, 5(1), 10–15. https://doi.org/10.1186/s40494-017-0123-8

Zapiola, A.G., Boselli, A., Menafoglio, A., et al., 2025. Hyper-spectral unmixing algorithms for remote compositional surface mapping: a review of the state of the art. Applied Computing and Geosciences, 28, 100297. https://doi.org/10.1016/j.acags.2025.100297