Towards the non-destructive analysis of multilayered samples: A novel XRF-VNIR-SWIR hyperspectral imaging system combined with multiblock data processing

Towards the non-destructive analysis of multilayered samples: A novel XRF-VNIR-SWIR hyperspectral imaging system combined with multiblock data processing

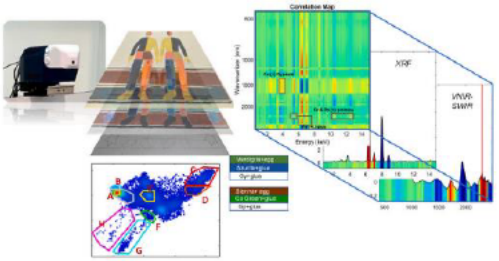
Emilio Catelli [a], Zelan Li [a], Giorgia Sciutto [a,*], Paolo Oliveri [b,**], Silvia Prati [a], Michele Occhipinti [c], Alessandro Tocchio [c], Roberto Alberti [c], Tommaso Frizzi [c], Cristina Malegori [b], Rocco Mazzeo [a]

[a] Department of Chemistry, University of Bologna-Ravenna Campus, via Guaccimanni, 42, 48121, Ravenna, Italy
[b] Department of Pharmacy, University of Genoa, via Cembrano, 4, 16148, Genoa, Italy
[c] XGLab SRL - Bruker Nano Analytics, Via Conte Rosso 23, I-20134 Milano, Italy

HIGHLIGHTS

- Novel whiskbroom hyperspectral scanner for high space-coherence chemical mapping.
- Simultaneous acquisition of XRF, VNIR and SWIR spectral ranges.
- Multiblock approach for the assessment of correlation between spectral ranges.
- Multivariate data processing for a thorough characterization of multi-layered samples.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Hyperspectral imaging
Chemometrics
Multiblock data processing
Paintings
VNIR SWIR spectroscopy
XRF analysis

ABSTRACT

The new challenge in the investigation of cultural heritage is the possibility to obtain stratigraphical information about the distribution of the different organic and inorganic components without sampling. In this paper recently commercialized analytical set-up, which is able to co-register VNIR, SWIR, and XRF spectral data simultaneously, is exploited in combination with an innovative multivariate and multiblock high-throughput data processing for the analysis of multilayered paintings. The instrument allows to obtain elemental and molecular information from superficial to subsurface layers across the investigated area. The chemometric strategy proved to be highly efficient in data reduction and for the extraction and integration of the most useful information coming from the three different spectroscopies, also filling the gap between data acquisition and data understanding through the combination of principal component analysis (PCA), brushing, correlation diagrams and maps (within and between spectral blocks) on the low-level fused. In particular, correlation diagrams and maps provide useful information for the reconstruction of a stratigraphic structure without the need to take any sample, thanks to the effective account for inter-correlation among data (variables), which is able to effectively characterize the possible combinations of components located in the same depth level. The highly innovative technology and the

* Corresponding author.
** Corresponding author.
E-mail addresses: giorgia.sciutto@unibo.it (G. Sciutto), paolo.oliveri@unige.it (P. Oliveri).

data processing strategy are applied for the multi-level characterization of a complex painting reproduction as an illustrative pilot study.

1. Introduction

To avoid damage on precious cultural heritage objects, the use of non-destructive techniques has been strongly encouraged. To this purpose, several research projects have been developed in the last few years to develop new advanced non-destructive and portable instrumentations [1]. Accordantly, a new challenge for the analytical chemistry in the investigation of artworks is the possibility to obtain stratigraphical information about the distribution of the different organic and inorganic components without sampling. Hyperspectral imaging (HSI) systems are today well established in a wide range of analytical chemistry applications, from medical diagnostics [2], food [3], and forensic sciences [4], to cultural heritage [5–8]. Thus, HSI technology has received great attention in many research fields due to the possibility of performing measurements in a low-invasive way, often without even touching the sample, together with the ability to record the information of large heterogeneous areas [5,9,10]. Indeed, HSI data are organized in a 3D data cube, in which the spatial information is allocated in the x and y dimension and the spectral information along the z dimension [11].

Hyperspectral imaging instrumentations span a wide portion of the electromagnetic spectrum, stretching from X-rays to the infrared region. In detail, commercial hyperspectral devices cover the subregions of X-ray (1–50 keV), UV (150–400 nm), Visible-NIR (VNIR) (400–1100 nm), SWIR (1100–2500 nm), MWIR (2500–5000 nm), and LWIR (8000–12000 nm). The instrumentations produce, as a result of the scanning, a 3D data cube or data block. Macro-X-ray fluorescence instrument (MA-XRF) is, among others, the most successful and largely applied system for non-invasive x-ray 2D surface scanning, achieving information about the elemental distribution across a region of interest [12,13]. Reflectance imaging spectroscopy technique operates in the UV–visible, near, short and mid infrared window, providing distribution of the chemical species by exploiting electronic and vibrational molecular transitions [13]. X-rays, near and short infrared waves, penetrate the analyzed object at a level depending on materials and their properties, giving information from superficial to subsurface features [14–17]. UV–Visible and mid-infrared waves offer mainly superficial information [14,17].

By combining spectral and spatial information in one single measurement, HSI allows to extract or create chemical images of the objects under study, displaying the distribution of the different chemical compounds, obtaining information also from inner layers when the employed technique exploits sources with penetrating radiation. As an example, the elemental distribution obtained with a macro XRF scanner allowed to discover original paintings hidden by overpainted layers [18,19].

It is worth to note that a critical point that limited the effective implementation and application of the hyperspectral systems, in different analytical sectors, is the complexity of the data processing. Each hyperspectral data block, in fact, is constituted by a massive amount of data that requires advanced computational and statistical methods to extract meaningful (chemical) information.

Moreover, in recent years, studies have shown the benefit of combining two or more hyperspectral imaging systems in complementary spectral regions. Delaney and co-workers proposed a single-point (whiskbroom) hyperspectral scanner to investigate paintings that co-collect VNIR-SWIR and XRF spectra [20]. Although very informative, the combination of several spectral blocks in one single 3D array matrix (through data fusion, DF) required additional processing steps and still presented a number of challenges. The most critical one is pixel alignment, required because the different instrumentations are characterized by different pixel sizes and spatial resolution [21,22]. Moreover, with the multiblock DF, the amount of data increased drastically, further emphasizing the need for computational strategies for data reduction and interpretation, also evaluating the relationship among the variables of the different blocks.

Within this scenario, Bruker Nano Analytics developed the first hyperspectral device that combines XRF with VNIR and SWIR reflectance in a whiskbroom scanner. This instrument, named IRIS, allows to simultaneously acquire X-ray fluorescence (XRF, 0–48 keV) data together with visible & near-infrared (VNIR, 380–1100 nm) and shortwave infrared reflectance spectroscopy (SWIR, 1100–2500 nm).

IRIS combines elemental and molecular information for the identification of inorganic and organic materials, from superficial to subsurface layers, across the total area of the painting. Indeed, SWIR and XRF spectroscopies are penetrating, reaching the subsurface painting layers. In particular, near-infrared radiation penetrates a few to one hundred microns [23], depending on several factors, including the energy of the source, the incident angle, and the material physical properties [24]. X-rays usually penetrate deeper, from tens to hundreds of micrometers [25–27], and collect information through all the layers down to the preparation one. Moreover, it is always worth considering that lighter elements are preferentially associated with the surface elements distribution, because their low-energy radiation is highly absorbed by the materials in the inner paint structure [28]. Data from these three spectral ranges (XRF, VNIR, and SWIR) are co-registered, thus inherently saved with high spatial coherence, reducing the time for post-processing and correlated analysis. Being co-aligned in the spatial dimension, the data blocks generated by IRIS are intrinsically multivariate, not only in the single spectral range but among the different spectral ranges, and thus a higher level of information in the identification of the compounds could be achieved by exploiting and interpreting the correlation among the variables. In the present study, IRIS was used for the first time in combination with an innovative data processing approach for the multi-level characterization of a complex painting reproduction in order to differentiate the contribution of the materials in different layers, by exploiting the penetration depth characteristic of the radiation employed. More in detail, an efficient multivariate strategy based on the combination of principal component analysis (PCA), brushing, correlation diagrams and maps (within and between spectral blocks) was purposedly developed. The chemometric strategy proved to be highly efficient in data reduction and for the extraction and integration of the most useful information coming from the three different spectroscopies, through effective data visualization and inter-correlation assessment.

To date, some data processing methods are gathered in the commercial software ENVI (L3Harris Geospatial Solutions, Inc.), designed mainly for remote sensing data processing [29]. Other methods are based on more flexible chemometric tools, such as principal component analysis (PCA), which can be easily combined with interactive tools such as brushing for an integrated interpretation of all PCA outcomes [30–32]. Recently developed methods such as t-distributed stochastic neighbor embedding (t-SNE) have been demonstrated to be useful, although their application was limited to classifying and mapping pigments [33].

2. Materials and methods

2.1. Laboratory painting mock-up

A complex painting reproduction was prepared at the Microchemistry and Microscopy Art Diagnostic Laboratory (M2ADL) of the University of Bologna, representing a freely inspired version of “Sportsmen” by Kazimir Malevich (ca. 1931, State Russian Museum,

Saint Petersburg). The heterogeneous and multi-layer reproduction was purposedly prepared to demonstrate the potential of the innovative scanner and multi-block HSI approach. It includes different areas which were painted using different pigments and binders. In more detail, a four-layer painted stratigraphy was prepared as follow: i) a wooden support, ii) a preparation layer made of gypsum and glue binder, iii) a first painting layer with glue binder, and iv) a second painting layer with egg binder. Ancient and modern pigments were used in both of the painting layers, applying the pigments in well-delimited sectors defining the two human figures and the background (Fig. 1 and Fig. S1). The pigments were selected from ancient and modern, bearing also in mind their intrinsic electronic, vibrational, and elemental signals detectable by VNIR, SWIR, and XRF spectroscopies, respectively. Pigment and binder were applied on the painting according to the scheme presented in Fig. 1a. The rendering of the preparation and painting layers is represented in Fig. 1b. The final view of the painting is presented in Fig. 1c. Pigments and binders were purchased from Zecchi (Florence, Italy) and Kremer Pigmente GmbH & Co. (Aichstetten, Germany).

2.2. Instrumentation

The scanner adopted for the measurements presented is IRIS (Bruker Nano Analytics), an analytical instrument developed in the framework of the project "MOBARTECH" [34], co-financed by Regione Lombardia (POR FESR 2014–2020).

IRIS (https://www.bruker.com/en/applications/academia-materials-science/art-conservation-archaeology/special-engineering.html) is a portable platform dedicated to synchronous and co-registered XRF and VNIR-SWIR hyperspectral measurement, with non-destructive and contactless approach to the sample. The XRF system is equipped with a rhodium-target X-ray tube and a nominal power of 10 W. The voltage range is between 10 and 50 kV, and the anode current range is between 5 and 200 μA. The X-ray fluorescence is revealed by a Peltier-cooled silicon drift detector (SDD) with CUBE technology, with an active area of 50 mm^2. The typical energy resolution for Mn-Kα radiation is < 140 eV with an input count rate of up to 500,000 cps. Three collimators with different size can be used (0.5, 1.0, 2.0 mm) to optimize the XRF measurement.

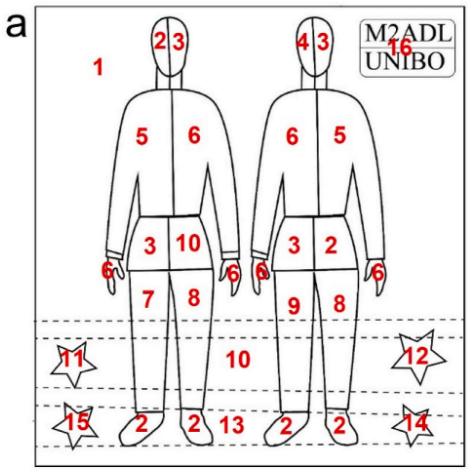
The hyperspectral technique performs better than 1.5 nm as spectral resolution in the most informative VNIR range (400–1100 nm), and better than 9.0 nm in the most informative SWIR region (1100–2500 nm). The intrinsic spatial resolution ranges between 0.6 and 0.8 mm depending on the wavelength.

Map analysis is performed with a whiskbroom scanning approach. The maximum extension of the map is 450 mm × 600 mm with a maximum speed of 42 mm/s.

The scanned area of the present experiment is 250 × 194 mm (249 × 193 pixels) with a collimator size of 1.0 mm and a spatial step along the x and y direction of 1 mm.

2.3. Multivariate image analysis approach

After IRIS hyperspectral data acquisition, the three raw data cubes, namely VNIR, SWIR, and XRF, have the same spatial dimension (same number of pixels) but different spectral dimensions (different number of variables). Since no particular spatial pre-processing was required, only



Area	Painted layer 1 + glue	Painted layer 2 + egg
1	Titanium white	Zinc white
2	Carbon black	Yellow Ochre
3	Lead white	Vermilion
4	Smalt	Sienna
5	Chrome yellow	Prussian blue
6	Sienna	Chrome yellow
7	Prussian blue	Lapis lazuli
8	Madder	Minium
9	Azurite	Lapis Lazuli
10	Malachite	Sienna
11	Malachite	Sienna
12	Cobalt green	Sienna
13	Sienna	Verdigris
14	Azurite	Verdigris
15	Smalt	Verdigris
16	Lead white	Zinc white

Fig. 1. Mock-up painting used in the research: (a) schematic representation and description of the pigments used in each area; (b) rendering of the preparatory and painting layers; (c) overview of the painting after its completion.

spectra pre-processing methods were applied. Firstly, VNIR and SWIR spectral variables were converted from reflectance into apparent absorbance (Log(1/R)). Each spectral range was edge-cut, considering the useful spectral region of 400–1087 nm for VNIR and 1087–2500 nm for SWIR. Successively, the baseline shifts and drifts observable in the SWIR region were corrected by applying a quadratic detrending. On the XRF data, a 4x binning along the spectral dimension was applied to improve the peak legibility and position, followed by a row-profile correction to remove global intensity (multiplicative) effects [35]. The XRF spectral range was also reduced, considering only the variables between 1 and 30 keV. After the above-mentioned operation, the spectral ranges consist of 1213 variables for the VNIR, 447 variables for the SWIR, and 307 variables for the XRF. To keep the VNIR and SWIR ranges with comparable spectral resolution, VNIR spectra were down-sampled taking one variable over three, reducing the total number of variables to 405. As the last step, the VNIR and SWIR data blocks were concatenated along the spectral dimension as they are contiguous in terms of the spectral variables and thus can be considered as a single block with variables included in the range from 400 to 2500 nm.

A low-level data fusion method (LLDF), which consisted in concatenating the VNIR-SWIR and XRF 3D data array along the spectral dimension after a separate block autoscaling, was chosen to merge the data.

The data processing method consisted of a multivariate strategy that consisted firstly of data exploration by principal component analysis (PCA) and secondly by the construction of correlation diagrams and maps, within and between the three spectral ranges examined. The approach was performed using a dedicated in-house software for image analysis, developed under MATLAB environment (v. 2019b, The MathWorks, Inc., Natick, MA, USA).

PCA was performed after mean-centering directly on the fused data cube, and results were firstly visualized as score images, where the score value of each pixel was encoded using a chromatic scale from blue (minimum score) to red (maximum score). Bidimensional scatter score plots of different pairs of PCs were then plotted using a density-based plotting method, here again using a chromatic scale from blue (minimum) to red (maximum) to individuate the low and the high-density pixel agglomerates and structures respectively. An interactive brushing approach was applied to these plots to understand the relations between the clusters in the score plot and the corresponding pixels in the original image domain. Such an approach proved very useful for identifying and studying pixel regions characterized by similar spectral features, which may also contain important information regarding similar chemical composition [30–32,36]. In detail, the brushing first allows the operator to manually select the cluster/s of interest in the score plot. The corresponding pixels are then automatically highlighted in the image. Moreover, the average spectral profiles of the selected pixels can be computed and visualized to enable chemical interpretation, together with PC loading analysis [32].

Correlation diagrams and maps, obtained using the VNIR-SWIR and XRF spectral blocks not submitted to block autoscaling, provided an easy visual highlighting of interrelations among variables in each spectral range and between the different ranges. Appropriate masks were created in the image to select the regions of interest (ROI), considering areas with similar chemical composition obtained through the previous PCA-brushing step). The ROIs were then submitted to the computation of the correlation coefficients.

Correlation diagrams and maps, obtained using the VNIR-SWIR and XRF spectral blocks not submitted to block autoscaling, provided an easy visual highlighting of interrelations among variables in each spectral range and between the different ranges. The basic idea of the correlation diagrams and maps is the computation of the correlation coefficient, a value that states the linear interdependence between pair of a selected variables. The correlation coefficient is thus computed for all the possible pairs of variables present in the dataset to build the correlation maps, either within the same spectral range or in different spectral ranges.

The values of the correlation coefficient may range between +1 and −1. A value of +1 or nearly close describes a behavior of the variables such that an increase in the first variable corresponds to an increase in the second variable (positive correlation). When the correlation coefficient value is −1, the opposite behavior (negative correlation) is verified. When the correlation coefficient value is close to zero, there is no correlation between the variables.

The correlation diagrams were calculated for the VNIR-SWIR and XRF spectral blocks separately. The graphical interface was programmed in order to select, on a representative spectral profile, a spectral variable of interest (wavelength or energy) on the abscissa. Subsequently, the correlation coefficient between a selected variable (indicated by a vertical red line that can be moved by the operator) and all the other variables in the range is graphically visualized using a chromatic scale from blue to red.

Visualization of the correlation between VNIR-SWIR and XRF blocks is obtained through correlation maps. In this case, correlation values are displayed in a 2D plot where the X-ray energies are on the abscissa, and the VNIR-SWIR wavelengths are on the ordinate. The overall workflow from data acquisition to data processing and visualization is presented in Fig. 2.

3. Results and discussion

The painting reconstruction described in the Materials and method section was analyzed with IRIS, and the data cube obtained was processed with a two-step multivariate procedure. In the first step, PCA and brushing were applied on the fused 3D array (where each pixel spectrum covers the extended range VNIR, SWIR, and XRF) to efficiently group or separate areas with similar chemical behavior. In the second step, a within- and between-block correlation was performed for a clear and solid representation of the correlation of critical variables. This step allowed to provide information about the multilayer composition of the painting reconstruction.

3.1. PCA and brushing

After PCA computation, score maps were visualized (Fig. S2) to identify the PCs yielding meaningful information (variance) to chemically characterize the different painted areas. Information that was retrieved from the most informative PC score maps (namely PC1, 3, 5and 6) was notable, allowing the identification of areas according to their chemical composition.

In the score plot PC1 vs. PC3, shown in Fig. 3a, it is possible to detect several score clusters – among which eight are clearly identifiable thanks to a higher point density – which correspond to the major areas of the painting in the image (Fig. 3b). Through the brushing procedure, these eight main clusters were manually selected in the score plot (A to H), and the related correspondence with the regions of the image was visualized. In the PC1 vs. PC6 score plot (Fig. 3c and d), it was possible to identify eight additional clusters (I to O), while a single but important cluster (P) was found in the PC1 vs. PC5 score plot, reported in Fig. 3f. It is worth mentioning that the cluster selection followed the criterion of analyzing the highest possible number of clusters for segmenting the image, which usually led to find the highest number of chemically different areas. For exploratory purposes, other clusters of scores were selected and visualized by brushing, also considering different combinations of PCs, but only the ones most informative for chemically describing painting areas were presented. By combining the information retained in the score images, brushing, and the corresponding averaged spectra per area, it was possible to identify painting details localized below the first painting layer.

An illustrative example of hidden details placed in the inner layer of the painting was represented by the star-shaped areas, which were visible in PC1 vs. PC3 and PC1 vs. PC6, and highlighted in clusters G, L,

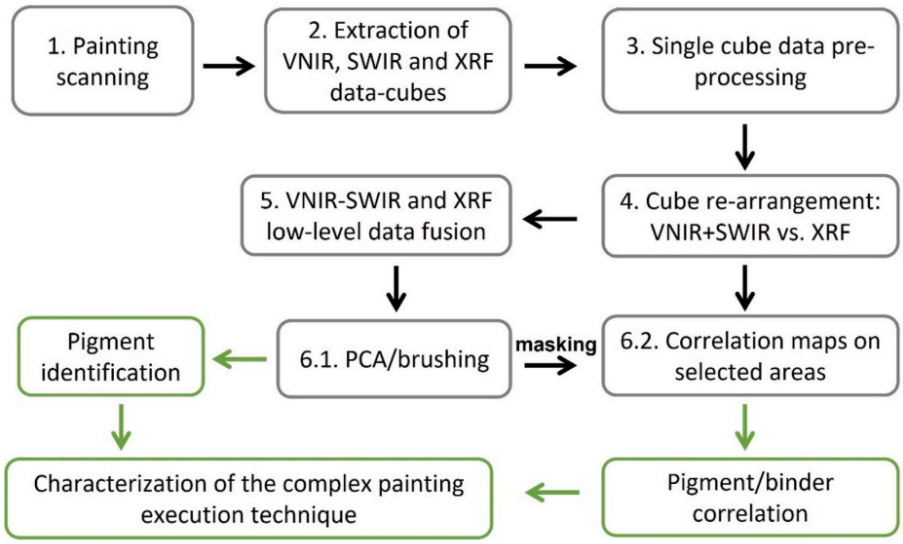


Fig. 2. Data collection, processing and visualization workflow.

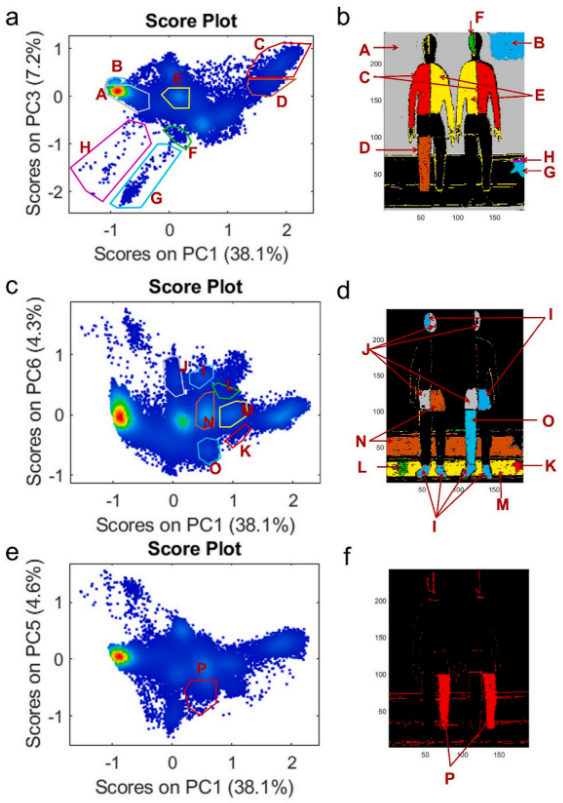


Fig. 3. (a) Score plot PC1 vs. PC3 with selected clusters (A to H); (b) score map with highlighted areas corresponding to A-H clusters; (c) score plot PC1 vs. PC6 with selected clusters (I to O); (d) score map with highlighted areas corresponding to I–O clusters; (e) score plot PC1 vs. PC5 with selected cluster (P); (f) score map with highlighted area corresponding to P cluster.

and K (Fig. 3 b,d)). The spectrum of the star defined by cluster G (Fig. 3b, Fig. 4(G)) revealed the presence of the characteristic transitions in the visible region at 580 nm (inflection point) and 850–900 nm (ligand field transition) of the iron-based pigment applied in the uppermost layer. Additionally, the cobalt green characteristic spectral shape in the 1300–1500 nm region [37] (Table S1) and Co, Ni, and Zn signals in the XRF region supported the identification of a cobalt-based pigment used for the creation of the star in the underneath layer. Similarly, the second star-shaped area (cluster K, Fig. 3d, Fig. 4(K)) showed the infrared bands of the pigment azurite (underneath layer) at 1489 nm [38] and verdigris (uppermost layer) between 2310 and 2260 nm [39]. The group of bands of verdigris were recognizable but not well resolved due to the overlap of azurite and lipidic bands. The presence of copper-based pigments was also confirmed by the XRF extracted spectrum (Fig. 4(K)), which allowed to detect only the Cu emission lines.

An additional star painted with smalt under the layer of verdigris, cluster L (Fig. 3d), was recognized thanks to the presence of a weak peak of cobalt in the X-ray region, which might suggest the presence of a cobalt-based pigment (Fig. 4(L)).

PCA investigation also allowed to identify areas in which, despite the lack of painted details present in the inner layer, the chemical compositions could be revealed clearly, and the presence of a two-layer structure might be hypothesized.

For example, the brown strip area (Area 10, cluster N in Fig. 3d) consisted of a layer of sienna earth and egg on top of a layer of malachite and glue below. The spectrum clearly showed the presence of the iron-based pigment, with the characteristic electronic transition [40] and the strong X-ray Fe signal (Fig. 5(N)). At the same time, the Cu X-ray signal, together with the marked shoulder at 2267 nm, suggested an additional contribution ascribable to malachite ($\nu 3\ CO_3^{2-}$) [39]. The visible brown color of the area suggested the presence of the malachite in the underneath layer, although the presence of a possible mixture of the two pigments may not be excluded.

In the white background area (Area 1, cluster A in Fig. 3b), it was possible to hypothesize the presence of two white pigments—titanium white and zinc white, due to the elements detected in XRF and the color appearance. However, these outcomes did not sufficiently allow to assert whether a mixture or multiple layers were applied to the painting. Similar results were obtained from another white area (Area 16, cluster B in Fig. 3b), characterized by the presence of lead white and zinc white, and from the area that presented vermilion in the uppermost layer and lead white underneath (Area 3, cluster J in Fig. 3d), for which could be not possible to exclude the use of lead white in mixture with vermillion in a single paint layer.

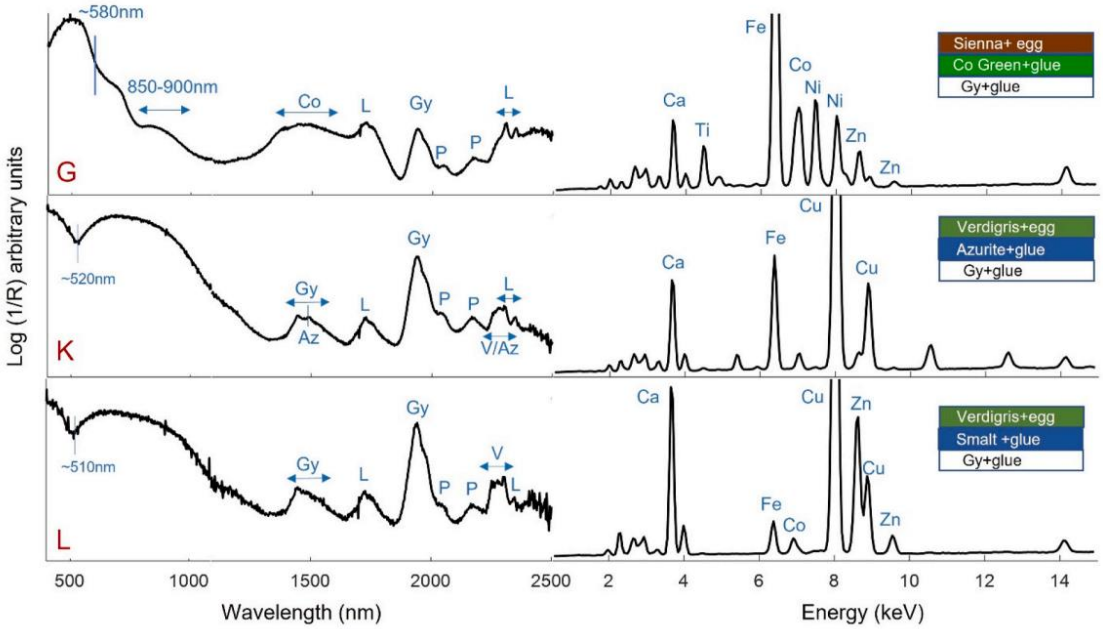


Fig. 4. Extracted average spectra and corresponding stratigraphy (top right) of the selected areas G, K and L. Abbreviations are used for the compounds and elements identified: Gy = gypsum, P = proteinaceous material, L = lipidic material, Az = azurite, V = verdigris, Fe = iron, Ca = calcium, Zn = zinc, Co = cobalt, Ni = nickel.

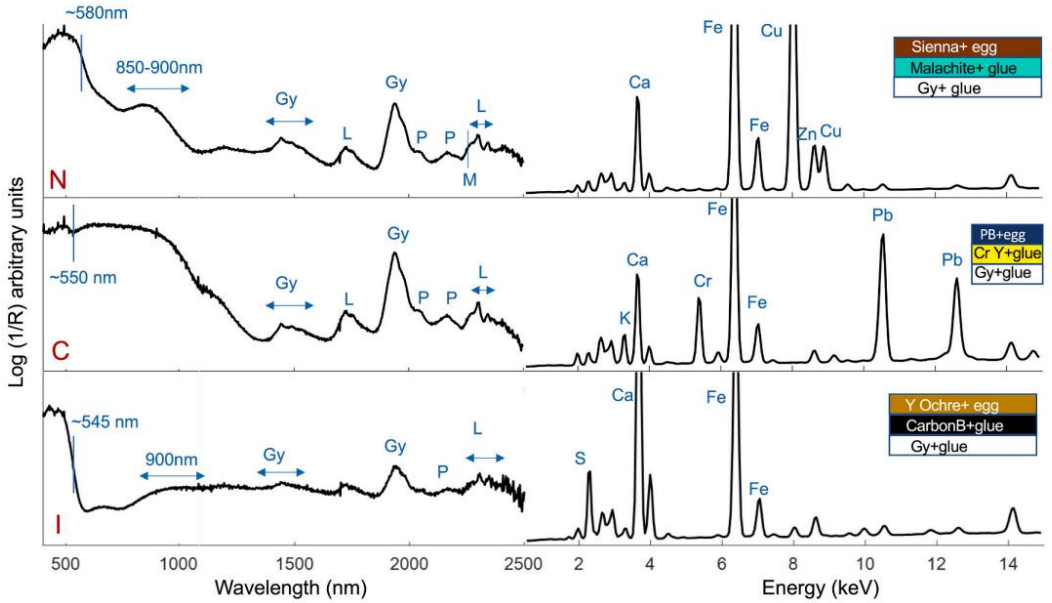


Fig. 5. Extracted average spectra and corresponding stratigraphy (top right) of the selected areas N,C, and I. Abbreviations are used for the compounds and elements identified: Gy = gypsum, P = proteinaceous material, L = lipidic material, M = malachite, Ca = calcium, Fe = iron, Cr = chromium, Zn = zinc, Pb = lead, K = potassium, S = sulfur, Cu = copper.

In the dark blue shirt's areas (Area 5, cluster C, Fig. 3b), made by the superimposition of a layer of Prussian blue and egg on a layer of chrome yellow and glue, the presence of Prussian blue could be inferred by the iron signal in the X-ray region together with the high absorption reflectance curve in the visible (minimum at 550 nm) (Fig. 5(C)) [41]. However, the presence of chrome and lead could suggest multiple compounds (i.e., lead carbonate, lead oxide, chrome oxide), and it was difficult to have a univocal attribution to chrome yellow.

Finally, in area 2 (cluster I in Fig. 3d), painted with yellow ochre and carbon black, the strong signal of Fe in the XRF spectrum indicated the presence of the iron-base pigment (Fig. 5(I)). Interestingly, the inflection point at 545 nm and the apparent absorbance maximum at 900 nm suggested that the yellow pigment applied on the uppermost layer could be yellow ochre [40]. Conversely, the carbon-based pigment is undetectable in the extended range considered: despite the very low reflectance values, the carbon-based pigment has no characteristic band in the

SWIR region, and the light element carbon is not detectable in the XRF spectrum.

Concerning the identification of the binders used overall in the painting, it was worth noting that, in the SWIR region, signals of proteinaceous materials at 2042 nm (ν(NH)+δ(NH)) and 2173 nm (2ν(CO) amide I + amide II) from both the first and the second paint layer were observed. Bands of lipids were also visible in the whole painting, due to the egg binder's fatty part at 2303 and 2347 nm ($\nu(CH_2)+\delta(CH_2)$) [42].

To deepen the evaluation of the PCA results, the different painted areas identified by PCA were submitted to a further investigation by using correlation diagrams. At this stage, the multivariate analysis was aimed at investigating more deeply the combined presence of components related to different types of pigments and/or present in different layers, exploiting the different penetration depths of the different spectroscopic techniques.

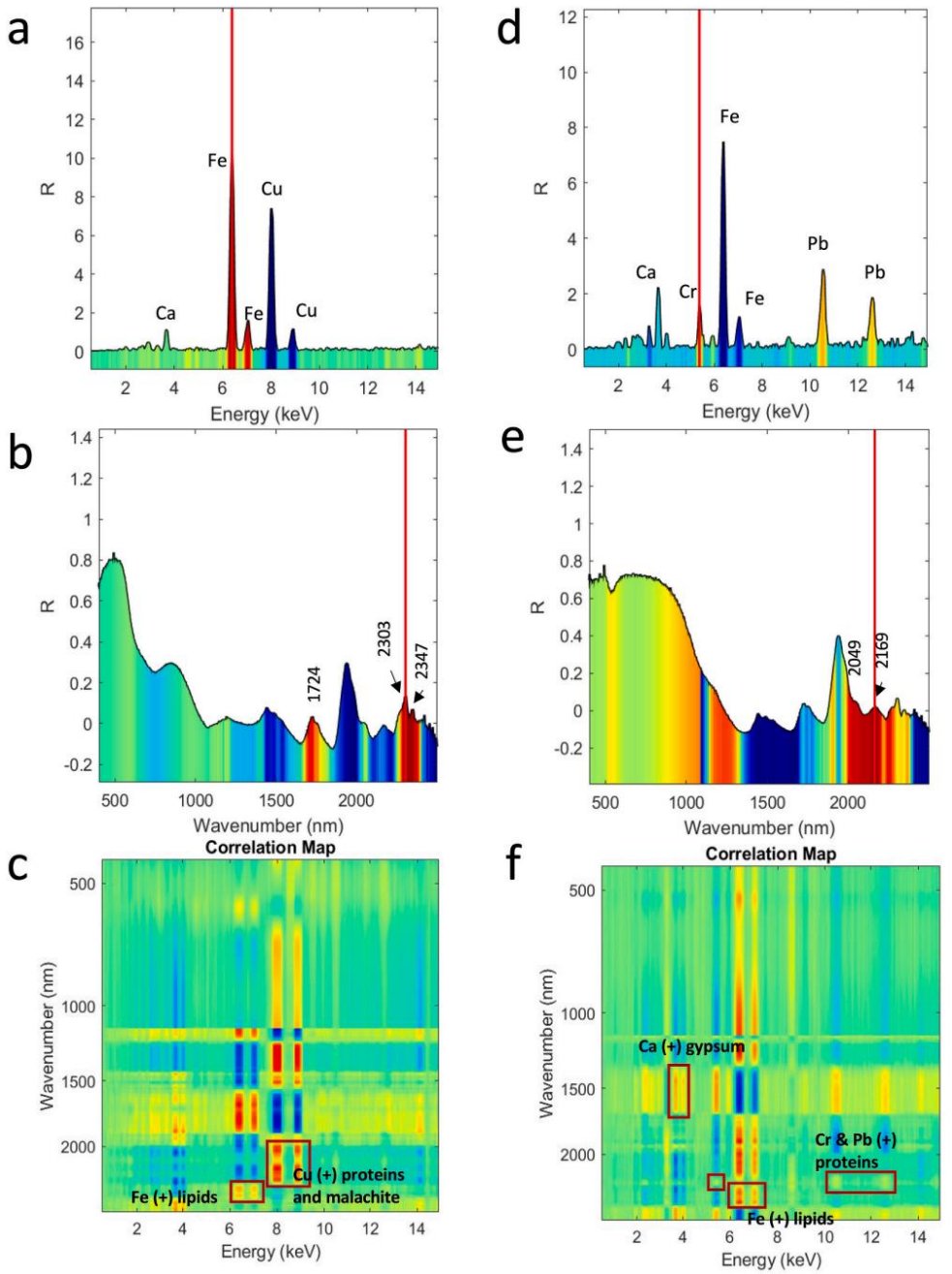


Fig. 6. Computed correlation for the molecular and elemental signals from the brown strip (cluster N) and the blue shirt (cluster C); (a)brown strip: within-block correlation of XRF; (b)brown strip: within-block correlation of VNIR-SWIR; (c)brown strip: between-block correlation for VNIR-SWIR vs. XRF; (d)blue shirt: within-block correlation of XRF; (e)blue shirt: within-block correlation of VNIR-SWIR; (f)blue shirt: between-block correlation for VNIR-SWIR vs. XRF. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Correlation diagrams

The results of within- and between-block spectral correlation of all the areas analyzed were summarized in Table S2. In this case, the VNIR-SWIR regions were considered as a single block, while the XRF data represented the second block.

Results obtained from the most representative areas, according to PCA outcomes, were reported below. As expected, the chemometric approach based on the evaluation of the correlation between and within spectral variables, provided additional information on the painting structure and composition. In Fig. 6 (a-b), the diagrams of the within-block correlation for the area associated with cluster N (Area 10) were shown. As described above, this brown-strip area was characterized by a first layer of malachite mixed with glue and an uppermost layer obtained with sienna pigment and egg.

The diagrams of the within-block correlation showed that the Cu peak, ascribable to the presence of malachite, presented a strong negative correlation (blue color) with Fe. The negative correlation means that when Cu values increased, Fe decreased with an inverse linear interdependence. This result suggested that these elements may not be ascribable to the same pigment. Since the area is homogenous, it could be supposed a mixture of two pigments or two pigments, possibly presented in two distinct paint layers. On the other hand, in the VNIR-SWIR spectrum, the fatty acids bands at 1724 nm, 2303 nm, and 2347 nm were positively correlated with each other (Fig. 6b) and not correlated to the proteinaceous-related bands at 2042 nm and 2173 nm, revealing the presence of the two distinct binders. Furthermore, in the related correlation map computed for this area (Fig. 6c), Cu at 8.0 keV was positively correlated with the proteinaceous signal at 2042 nm and 2173 nm, while it was negatively correlated to fatty materials, as the color at the corresponding coordinates appeared blue. Moreover, the SWIR shoulder at 2271 nm, ascribable to the second overtone of the CO_3^{2-} antisymmetic stretching of carbonate in malachite [39], showed a positive within-block correlation with the proteinaceous bands, confirming the presence of malachite and of a proteinaceous binder, mixed in the same layer. A positive correlation between Fe and the fatty materials was observed, confirming their simultaneous presence in the same layer.

Thus, according to the results and based on the visual appearance of the layer, it is possible to hypothesize the presence in this cluster of two superimposed layers, the inner one made of malachite and a proteinaceous binder, the superficial one composed of an iron-based pigment applied with a lipidic material.

Concerning the area of the blue shirts (Area 5, cluster C), the XRF correlation diagram (Fig. 6d) showed chromium and lead positively correlated (red color) while iron is negatively correlated with both (blue color), suggesting the presence of two pigments, one based on iron and one based on chromium and lead, possibly lead chromate. Also, in this case, the absence of correlation between proteinaceous and lipidic bands suggested that the area has two distinct binders (Fig. 6e). In the correlation map (Fig. 6f), it was possible to observe a positive correlation between the combination bands of the proteinaceous material (2049 and 2169 nm) and chromium and lead elements, corresponding to the actual composition of the inner painting layer. Furthermore, the combination bands of the lipidic binder (2304 and 2346 nm) were correlated to the iron peak of Prussian blue, according to the composition of the outer layer. A positive correlation was also revealed between the near-infrared bands of gypsum (OH stretching overtone) and the calcium XRF peak (Fig. 6f). This second example further proved the possibility of distinctly identifying the correspondence between pigments and binders, thus directing data interpretation in a more accurate way to identify pigment-binder mixtures in distinct layers. Also, in this case, the visual appearance of the painting may suggest that the external layer is composed of Prussian blue applied with a lipidic material, with the lead and chrome-based pigment coming from an inner layer applied with a proteinaceous binder.

Similar to the previous examples, it was possible to hypothesize two painting layers in the white background areas (Fig. 1 areas 1 and 16) since the two elements were not correlated, while the titanium was positively correlated with a proteinaceous binder while zinc was with the lipidic binder (Table S2). In this case, being both the two layer white, it is not possible to state which layer is the external one. As further development, the combination with other non-destructive superficial techniques may support the correct identification of the stratigraphic distribution.

The great advantage of the methodology proposed here is the possibility to progressively isolate the areas with similar chemical composition and use effective tools to thoroughly characterize the materials inside each area. In the elemental univariate XRF analysis (Fig. S3), for example, the map of a single element usually represents the overall distribution of the element in the painting, leading to some difficulty and uncertainty in the pigment characterization, especially when an element is shared between two or more pigments. In our method, we use PCA-brushing first to isolate the chemically similar areas. Secondly, from each area, both the distribution and correlation of the element with respect to the other elements can be provided, supporting the compounds' characterization. The VNIR-SWIR correlation follows the same structure. In the end, the between-block correlation map of VNIR-SWIR and XRF provides, for the same area, the pigment-binder relationship with additional possibilities for the allocation of multiple layers.

4. Conclusions

In this research work, a state-of-the-art hyperspectral system was applied for the first time, combining three different spectroscopies with high spatial coherence, together with an advanced multivariate and multi-block data processing approach. The study demonstrated the possibility of addressing, with such a system, new urgent issues of the analytical chemistry applied to the characterization of the composition of multi-layered objects without the need of performing micro sampling. To this purpose, the research considered the fundamental role that data processing is assuming in various areas, filling the gap between the big data acquired and the possibility of extracting useful information from them. The results obtained demonstrated the ability of the proposed method to provide highly characterizing chemical information in a non-invasive way, exploiting the capabilities of an unsupervised chemometric approach. In particular, multivariate exploratory analysis led to identifing and mapping composition variability, while the evaluation of the within- and between-block correlation, revealed the relationship among pigments and binders, which led to robust hypotheses on paint stratigraphy and execution techniques. The possibility – offered by the method described in the present study – of achieving stratigraphical information on the composition of multi-layered objects in a non-destructive way will be of great interest in many application sectors, including the industrial, food, forensic, and medical fields. Moreover, in order to improve the performance of the system in the non-destructive reconstruction of the multilayers composition, further research is in progress to combine IRIS data with other data arrays obtained with other hyperspectral systems which provide superficial information (i.e., macro FTIR scanner).

CRediT authorship contribution statement

Emilio Catelli: Formal analysis, Data curation, Writing – original draft, Visualization. **Zelan Li:** Formal analysis, Writing – original draft. **Giorgia Sciutto:** Conceptualization, Methodology, Data curation, Writing – review & editing, Supervision, Visualization. **Paolo Oliveri:** Conceptualization, Methodology, Writing – review & editing, Supervision, Software. **Silvia Prati:** Writing – review & editing, Visualization, Validation. **Michele Occhipinti:** Investigation, Writing – original draft. **Alessandro Tocchio:** Writing – review & editing, Validation. **Roberto Alberti:** Investigation. **Tommaso Frizzi:** Writing – review & editing, Validation. **Cristina Malegori:** Formal analysis. **Rocco Mazzeo:**

Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

Financial support provided by Università degli Studi di Genova (Research Project Curiosity Driven 2020: "3Depth – From 2D to 3D hyperspectral imaging exploiting the penetration depth of near-infrared radiation", CUP: D34G20000100005) is gratefully acknowledged.

We are thankful to Dr. Diego Quintero for the thorough preparation of the mock-up sample used in this study. The support provided by China Scholarship Council (CSC) during the Ph.D. research of Zelan Li at the University of Bologna is acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2022.340710.

References

[1] https://www.iperionhs.eu/. (Accessed 1 September 2022).
[2] G. Lu, B. Fei, Medical hyperspectral imaging: a review, J. Biomed. Opt. 19 (1) (2014), 010901.
[3] G. Özdoğan, X. Lin, D.W. Sun, Rapid and noninvasive sensory analyses of food products by hyperspectral imaging: recent application developments, Trends Food Sci. Technol. 111 (2021) 151–165.
[4] G.J. Edelman, E. Gaston, T.G. Van Leeuwen, P.J. Cullen, M.C.G. Aalders, Hyperspectral imaging for non-contact analysis of forensic traces, Forensic Sci. Int. 223 (1–3) (2012) 28–39.
[5] M. Kubik, Hyperspectral Imaging: a New Technique for the Non-invasive Study of Artworks in Physical Techniques in the Study of Art, Archaeology and Cultural Heritage, Elsevier, 2007, pp. 199–259.
[6] C. Fischer, I. Kakoulli, Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications, Stud. Conserv. 51 (2006) 3–16.
[7] H. Liang, Advances in multispectral and hyperspectral imaging for archaeology and art conservation 106 (2) (2012) 309–323.
[8] J. Striova, A. Dal Fovo, R. Fontana, Reflectance imaging spectroscopy in heritage science, La Rivista del Nuovo Cimento (2020) 1–52.
[9] G.J. Edelman, E. Gaston, T.G. Van Leeuwen, P.J. Cullen, M.C. Aalders, Hyperspectral imaging for non-contact analysis of forensic traces, Forensic Sci. Int. 223 (1–3) (2012) 28–39.
[10] M. Halicek, H. Fabelo, S. Ortega, G.M. Callico, B. Fei, In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer, Cancers 11 (6) (2019) 756.
[11] M. Vidal, J.M. Amigo, Pre-processing of hyperspectral images. Essential steps before image analysis, Chemometr. Intell. Lab. Syst. 117 (2012) 138–148.
[12] K. Langstraat, A. Knijnenberg, G. Edelman, L. Van De Merwe, A. van Loon, J. Dik, A. van Asten, Large area imaging of forensic evidence with MA-XRF, Sci. Rep. 7 (1) (2017) 1–11.
[13] M. Alfeld, L. De Viguerie, Recent developments in spectroscopic imaging techniques for historical paintings-a review, Spectrochim. Acta B: Atom. Spectros 136 (2017) 81–105.
[14] M.V. Padalkar, N. Pleshko, Wavelength-dependent penetration depth of near infrared radiation into cartilage, Analyst 140 (7) (2015) 2093–2100.
[15] M. Huang, M.S. Kim, K. Chao, J. Qin, C. Mo, C. Esquerre, S. Delwiche, Q. Zhu, Penetration depth measurement of near-infrared hyperspectral imaging light for milk powder, Sensors 16 (4) (2016) 441.
[16] C.G. Schroer, P. Cloetens, M. Rivers, A. Snigirev, A. Takeuchi, W. Yun, High-resolution 3D imaging microscopy using hard X-rays, MRS Bull. 29 (3) (2004) 157–165.
[17] J.Q. Brown, K. Vishwanath, G.M. Palmer, N. Ramanujam, Advances in quantitative UV–visible spectroscopy for clinical and pre-clinical application in cancer, Curr. Opin. Biotechnol. 20 (1) (2009) 119–131.
[18] S. Mosca, T. Frizzi, M. Pontone, R. Alberti, L. Bombelli, V. Capogrosso, A. Nevin, G. Valentini, D. Comelli, Identification of pigments in different layers of illuminated manuscripts by X-ray fluorescence mapping and Raman spectroscopy, Microchem. J. 124 (2016) 775–784.
[19] G. Van der Snickt, S. Legrand, I. Slama, E. Van Zuien, G. Gruber, K. Van der Stighelen, L. Klaassen, E. Oberthaler, K. Janssens, In situ macro X-ray fluorescence (MA-XRF) scanning as a non-invasive tool to probe for subsurface modifications in paintings by P.P. Rubens, Microchem. J. 138 (2018) 238–245.
[20] J.K. Delaney, D.M. Conover, K.A. Dooley, L. Glinsman, K. Janssens, M. Loew, Integrated X-ray fluorescence and diffuse visible-to-near-infrared reflectance scanner for standoff elemental and molecular spectroscopic imaging of paints and works on paper, Herit. Sci. 6 (1) (2018) 1–12.
[21] M. Alfeld, S. Pedetti, P. Martinez, P. Walter, Joint data treatment for Vis–NIR reflectance imaging spectroscopy and XRF imaging acquired in the Theban Necropolis in Egypt by data fusion and t-SNE, C. R. Phys. 19 (7) (2018) 625–635.
[22] S. Kahraman, R. Bacher, A comprehensive review of hyperspectral data fusion with lidar and sar data, Annu. Rev. Control 51 (2021) 236–253.
[23] M. Longoni, B. Genova, A. Marzanni, D. Melfi, C. Beccaria, S. Bruni, FT-NIR spectroscopy for the non-invasive study of binders and multi-layered structures in ancient paintings: artworks of the lombard renaissance as case studies, Sensors 22 (5) (2022) 2052.
[24] P. Geladi, Are pixels sample cells? Hyperspectral diffuse near infrared imaging experiments with pinholes, J. Near Infrared Spectrosc. 16 (3) (2008) 357–363.
[25] L. De Viguerie, V.A. Sole, P. Walter, Multilayers quantitative X-ray fluorescence analysis applied to easel paintings, Anal. Bioanal. Chem. 395 (7) (2009) 2015–2020.
[26] M. Eveno, B. Moignard, J. Castaing, Portable apparatus for in situ X-ray diffraction and fluorescence analyses of artworks, Microsc. Microanal. 17 (5) (2011) 667–673.
[27] M. Alfeld, W. De Nolf, S. Cagno, K. Appel, D.P. Siddons, A. Kuczewski, K. Janssens, J. Dik, K. Trentelman, M. Walton, A. Sartorius, Revealing hidden paint layers in oil paintings by means of scanning macro-XRF: a mock-up study based on Rembrandt's "An old man in military costume,", J. Anal. At. Spectrom. 28 (1) (2013) 40–51.
[28] P. Ricciardi, S. Legrand, G. Bertolotti, K. Janssens, Macro X-ray fluorescence (MA-XRF) scanning of illuminated manuscript fragments: potentialities and challenges, Microchem. J. 124 (2016) 785–791.
[29] https://www.l3harrisgeospatial.com/Software-Technology/ENVI. (Accessed 1 September 2022).
[30] P. Geladi, H. Grahn, M. Manley, In Raman, Infrared, and Near- Infrared Chemical Imaging, Wiley Ed., Hoboken, 2010, pp. 93–107.
[31] G. Sciutto, P. Oliveri, S. Prati, M. Quaranta, S. Bersani, R. Mazzeo, An advanced multivariate approach for processing X-ray fluorescence spectral and hyperspectral data from non-invasive in situ analyses on painted surfaces, Anal. Chim. Acta 752 (2012) 30–38.
[32] G. Sciutto, P. Oliveri, S. Prati, E. Catelli, I. Bonacini, R. Mazzeo, A multivariate methodological workflow for the analysis of FTIR chemical mapping applied on historic paint stratigraphies, Int. J. Anal. Chem. 2017 (2017) 1–12.
[33] E. Pouyet, N. Rohani, A.K. Katsaggelos, O. Cossairt, M. Walton, Innovative data reduction and visualization strategy for hyperspectral imaging datasets using t-SNE approach, Pure Appl. Chem. 90 (3) (2018) 493–506.
[34] https://mobartech.unimib.it/en/. (Accessed 1 September 2022).
[35] P. Oliveri, C. Malegori, R. Simonetti, M. Casale, The impact of signal pre-processing on the final interpretation of analytical outcomes–A tutorial, Anal. Chim. Acta 1058 (2019) 9–17.
[36] J.M. Prats-Montalbán, A. de Juan, A. Ferrer, Multivariate image analysis: a review with applications, Chemometr. Intell. Lab. 107 (1) (2011) 1–23.
[37] L. Boselli, Non-invasive spectroscopic study of 19th century artists' materials, Università degli Studi di Ferrara (Ph.D. Thesis), 2010.
[38] K.A. Dooley, S. Lomax, J.G. Zeibel, C. Miliani, P. Ricciardi, A. Hoenigswald, M. Loew, J.K. Delaney, Mapping of egg yolk and animal skin glue paint binders in Early Renaissance paintings using near infrared reflectance imaging spectroscopy, Analyst 138 (17) (2013) 4838–4848.
[39] D. Buti, F. Rosi, B.G. Brunetti, C. Miliani, In-situ identification of copper-based green pigments on paintings and manuscripts by reflection FTIR, Anal. Bioanal. Chem. 405 (8) (2013) 2699–2711.
[40] M. Aceto, A. Agostino, G. Fenoglio, A. Idone, M. Gulmini, M. Picollo, P. Ricciardi, J.K. Delaney, Characterisation of colourants on illuminated manuscripts by portable fibre optic UV-visible-NIR reflectance spectrophotometry, Anal. Methods 6 (5) (2014) 1488–1500.
[41] J.K. Delaney, J.G. Zeibel, M. Thoury, R.O.Y. Littleton, M. Palmer, K.M. Morales, E. R. de La Rie, A.N.N. Hoenigswald, Visible and infrared imaging spectroscopy of Picasso's Harlequin musician: mapping and identification of artist materials in situ, Appl. Spectrosc. 64 (6) (2010) 584–594.
[42] M. Vagnini, C. Miliani, L. Cartechini, P. Rocchi, B.G. Brunetti, A. Sgamellotti, FT-NIR spectroscopy for non-invasive identification of natural polymers and resins in easel paintings, Anal. Bioanal. Chem. 395 (7) (2009) 2107–2118.