Journal of Visualized Experiments

Applying Hyperspectral Reflectance Imaging to investigate the palettes and the techniques of painters

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE62202R3
Full Title:	Applying Hyperspectral Reflectance Imaging to investigate the palettes and the techniques of painters
Corresponding Author:	Michele Caccia University of Milan–Bicocca: Universita degli Studi di Milano-Bicocca Milano, ITALY
Corresponding Author's Institution:	University of Milan–Bicocca: Universita degli Studi di Milano-Bicocca
Corresponding Author E-Mail:	michele.caccia@unimib.it
Order of Authors:	Anna Galli
	Michele Caccia
	Simone Caglio
	Matteo Interlenghi
	Isabella Castiglioni
	Marco Martini
Additional Information:	
Question	**Response**
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$2,400)
Please specify the section of the submitted manuscript.	Chemistry
Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations.	Milan, Italy
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please provide any comments to the journal here.	
Please indicate whether this article will be Standard Access or Open Access.	Standard Access ($1400)

TITLE:

Applying Hyperspectral Reflectance Imaging to Investigate the Palettes and the Techniques of Painters

AUTHORS AND AFFILIATIONS:

Michele Caccia[1], Simone Caglio[1], Anna Galli[1,2], Matteo Interlenghi[2], Isabella Castiglioni[2,3], Marco Martini[1]

[1]Dipartimento di Scienza dei Materiali, Università degli Studi di Milano-Bicocca, Milano, Italy
[2]CNR-IBFM, Segrate, Milano, Italy
[3]Dipartimento di Fisica, Università degli Studi di Milano-Bicocca, Milano, Italy

Corresponding author:
Michele Caccia (michele.caccia@unimib.it)

Email addresses of co-authors:
Michele Caccia (michele.caccia@unimib.it)
Simone Caglio (simone.caglio@unimib.it)
Anna Galli (anna.galli@unimib.it)
Matteo Interlenghi (m.interlenghi@gmail.com)
Isabella Castiglioni (castiglioni.isabella@gmail.com)
Marco Martini (m.martini@unimib.it)

Keywords:

Hyperspectral Reflectance Imaging, Reflectance Spectroscopy, Fibre Optics Reflectance Spectroscopy, Spectral Angle Mapper, data manipulation, custom-adjustable analysis, paintings, pigments

Summary:

Hyperspectral Reflectance Imaging hypercubes include remarkable information into a large amount of data. Therefore, the request for automated protocols to manage and study the datasets is widely justified. The combination of Spectral Angle Mapper, data manipulation, and a user-adjustable analysis method constitutes a key-turn for exploring the experimental results.

Abstract:

Reflectance Spectroscopy (RS) and Fiber Optics Reflectance Spectroscopy (FORS) are well-established techniques for the investigation of works of art with particular attention to paintings. Most modern museums put at the disposal of their research groups portable equipment that, together with the intrinsic non-invasiveness of RS and FORS, makes possible the *in situ* collection of reflectance spectra from the surface of artefacts. The comparison, performed by experts in pigments and painting materials, of the experimental data with databases of reference spectra drives the characterization of the palettes and of the techniques used by the artists. However, this approach requires specific skills, and the results are time consuming especially if the number of the spectra to be investigated becomes large as is the case of Hyperspectral Reflectance Imaging (HRI) datasets. The HRI experimental setups are multi-dimensional cameras that associate the spectral information, given by the

reflectance spectra, with the spatial localization of the spectra over the painted surface. The resulting datasets are 3D-cubes (called hypercubes or data-cubes) where the first two dimensions locate the spectrum over the painting and the third is the spectrum itself (i.e., the reflectance of that point of the painted surface versus the wavelength in the operative range of the detector). The capability of the detector to simultaneously collect a great number of spectra (typically much more than 10,000 for each hypercube) makes the HRI datasets large reservoirs of information and justifies the need for the development of robust and, possibly, automated protocols to analyze the data. After the description of the procedure designed for the data acquisition, we present an analysis method that systematically exploits the potential of the hypercubes. Based on Spectral Angle Mapper (SAM) and on the manipulation of the collected spectra, the algorithm handles and analyzes thousands of spectra while at the same time supporting the user to unveil the features of the samples under investigation. The power of the approach is illustrated by applying it to *Quarto Stato*, the iconic masterpiece by Giuseppe Pellizza da Volpedo, held in the Museo del Novecento in Milan (Italy).

Introduction:

Reflectance Spectroscopy (RS) and Fiber Optics Reflectance Spectroscopy (FORS) are based on the detection of the light reflected by surfaces once illuminated by a light source, typically a tungsten-halogen lamp. The output of the acquisition system is spectra where the reflectance is monitored as a function of the wavelength in a range that depends on the characteristics of the employed experimental setup[1–3]. Introduced during the last four decades[4,5], RS and FORS are typically used in combination with X-ray fluorescence and other spectroscopies to describe the materials and the techniques used by artists to realize their masterpieces[6–9]. The study of the reflectance spectra is usually performed by comparing the data from the sample with a group of reference spectra selected by the user in personal or public databases. Once users identify references that comply with the realization period of the sample and with the *modus operandi* of the artist, users recognize the main features of the reflectance spectra (i.e., transition, absorption, and reflection bands[1,2,10,11]). Then, users supported by other techniques[6–8] distinguish the pigments have been used in the paintings and identify the slight differences between the references and the experimental spectra[7,9].

In most cases, the experimental datasets are composed of a few spectra, collected from areas chosen by art experts and assumed to be significant for the characterization of the painting[6,12,13]. Despite the skills and the experience of the user, a few spectra cannot fully exhaust the characteristics of the whole painted surface. Moreover, the result of the analysis will always be strongly dependent on the expertise of the performer. In this scenario, Hyperspectral Reflectance Imaging (HRI[3,14,15]) could be a useful resource. Instead of a few isolated spectra, the experimental setups return the reflectance properties of extended portions or even of the whole artefact under investigation[16]. The two main advantages with respect to the acquisition of the isolated spectra are evident. On one hand, the availability of the spatial distribution of the reflectance properties allows the identification of areas that hide interesting features, even though it may not seem peculiar[17]. On the other hand, the hypercubes guarantee a number of spectra high enough to enable the statistical analysis of the data. These facts support the comprehension of the distribution of pigments within the painted surface[18,19].

With HRI, the comparison of the experimental data with the references could be hard to handle[15]. A typical detector returns hypercubes of at least 256 x 256 spectra. This would require the user to evaluate more than 65,000 reflectance spectra against each reference, a task almost impossible to be carried out manually in a reasonable time. Therefore, the request for robust and, possibly, automated protocols to manage and analyze HRI datasets is more than justified[15,17]. The proposed method answers this need by handling the whole analytic procedure with the minimum involvement and the maximum flexibility.

An algorithm comprising a set of home-made codes (**Table of Materials**) reads, manages, and organizes the files returned by the experimental setup. It allows the fine selection of the portions of the Fields of View (FOVs, the area of the painting monitored by a single hypercube) to be studied and performs the analysis of the data based on the Spectral Angle Mapper (SAM) method[20,21] and on the manipulation of the original spectra. SAM returns false color gray-scale images called similarity maps. The values of the pixels of these maps correspond to the spectral angles (i.e., the angles between the spectra stored in the hypercubes) and the so-called End Members (EMs, a group of reference spectra that should describe the features of the surface monitored by the hypercubes)[22]. In the case of RS applied to paintings, the EMs are the reflectance spectra of pigments that should match the palette of the Master. They are chosen based on the available information about the artist, the realization period of the painting, and the expertise of the user. Therefore, the output of the SAM is a set of maps that describes the spatial distributions of these pigments over the painting surface and that supports the user to infer the materials used by the artist and their organization in the artefact. The algorithm offers the possibility of employing all kind of references independently from their origin. The references can be specific spectra selected within the hypercubes, come from databases, be acquired by a different instrument on a different surface (such as samples of pigments or the palette of the artist, for instance), or be obtained employing any kind of reflectance spectroscopy, FORS included.

SAM has been preferred among the available classification methods (e.g., the main classification methods of Richard[23]) because it is effective for easily characterizing pigments to be understood and implemented. Instead, the idea of developing a home-made protocol rather than adopting one of the many tools freely available on the net[24,25] relies on a practical consideration. Despite the effectiveness and scientific foundation of the existing GUIs and software, a single tool hardly satisfies all the needs of the user. There could be an Input/Output (I/O) issue because a tool does not manage the file containing the raw data. There could be an issue regarding the analysis of the data because another tool does not provide the desired approach. There could be a limitation in the handling of the data because the simultaneous analysis of multiple datasets is not supported. In any case, a perfect tool does not exist. Each method must be adjusted to the data or vice versa. Therefore, the development of a home-made protocol has been preferred.

The presented approach offers neither a complete set of analytical approaches (see, for comparison, the tool proposed by Mobaraki and Amigo[24]) nor an easy-to-manage user-interface (see, for comparison, the software employed by Zhu and co-workers[25]), but, in exchange, it focuses on a still underrated aspect of hyperspectral data analysis: the opportunity to manipulate the detected spectra. The power of the approach is illustrated by applying it to the painting *Quarto Stato* by Giuseppe Pellizza da Volpedo (**Figure 1**), an oil on canvas held in the Museo del Novecento in Milan, Italy. Note that, since the approach requires running home-made codes, the developer arbitrarily chose the names of the codes and both the input and output variables used in the description of the protocol. The user can change the names of the variables as they prefer but they must provide (i) the input variables within brackets and eventually separated by comma and (ii) the output variables within square brackets and eventually separated by a white space. On the contrary the names of the codes cannot be altered.

Protocol:

1. Set the spatial resolution of the hypercubes

1.1. Perform a preliminary inspection of the painted surface (**Figure 1**) supported by art experts to identify the main features of the painting.

1.1.1. Recognize the pictorial techniques employed by the artist to create the painting.

1.1.2. Identify the different brush strokes of paint on the canvas.

1.1.3. Estimate, qualitatively, the characteristics of the brush strokes with particular attention to their size.

1.2. Mimic the pictorial technique used by the artist by creating *ad-hoc* test samples where the brush strokes show characteristics similar to those applied by the artist.

NOTE: Pellizza da Volpedo was a Divisionist painter. A restorer was asked to prepare some test samples that qualitatively reproduce the brush strokes of the canvas of interest (**Figure 2, column A**).

1.3. Set the distance between the surface under investigation and the acquisition equipment.

NOTE: The distance determines the spatial resolution of the hypercubes[26] and therefore the possibility to distinguish the brush strokes on the images and SAM maps of the painted surface.

1.3.1. Evaluate the distance between the surface of the sample and the acquisition equipment based on the characteristics of the hyperspectral camera[26] (**Table of Materials**) and on the size of the brush strokes drawn in the test samples.

1.3.2. Put the acquisition stage and the hyperspectral camera at the distance evaluated in the previous step. Arrange the test samples on the stage and ensure uniform illumination of the surface of the samples.

1.3.3. Perform a white calibration using the white standard reference supplied with the hyperspectral camera. Acquire the hypercubes.

NOTE: For each FOV, the hyperspectral camera returns both raw and calibrated images. The latter have been used for the analysis.

1.3.4. Download the files returned by the instrument and save them in a dedicated folder.

1.4. Check whether the spatial resolution of the hypercubes can distinguish the different brush strokes on the images and SAM maps of the painted surface.

1.4.1. Inspect the RGB pictures returned by the hyperspectral camera to ensure that the brush strokes used to realize the test samples can be recognized (**Figure 2, column A**). If so, move to the next steps; otherwise go back to step 1.3.1 and restart.

1.4.2. List the files containing the hyperspectral data and their RGB images of the FOVs by running the data reading code, **HS_FileLister**. Type the following command line (semicolon included) in the terminal window of the language used to develop the codes (**Table of Materials**) and press **Enter** to run the code:

[**HS_DataList HS_ImageList] = HS_FileLister;**

1.4.2.1. No input is required and there are two outputs: the list of the files containing the hypercubes, **HS_DataList**, and the list of the images returned by the hyperspectral camera, **HS_ImageList**.

NOTE: The size of each hypercube is 512 x 512 x 204 voxels where 204 is the number of channels used to monitor the reflectance signal. The channels span the wavelength range between 400 and 1,000 nm with a spectral resolution of 7 nm at FWHM[26].

1.4.3. Define the 3D portion of the hypercubes that must be analyzed by running the cropping code, **HS_Crop_png**. Define the desired portion of each data-cube by selecting an area over an interactive window that shows the 2D, RGB image of the FOV monitored by each hypercube. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

[**HS_ImageList] = HS_Crop_png(HS_ImageList);**

1.4.3.1. There is one input (the list of the images returned by the hyperspectral camera, **HS_ImageList**) and one output (the input list added with the spatial coordinates to eventually crop the hypercubes).

1.4.4. Apply the D65 illuminant and 1931 observer from CIE (International Commission on Illumination) standards to the hypercubes to retrieve the RGB images of the FOV(s) from the reflectance spectra by running the re-building code, **HS_RGB_rebuild**. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

[HS_ImageList] = HS_RGB_rebuild(HS_ImageList, HS_DataList);

1.4.4.1. There are two inputs (the list containing the images returned by the hyperspectral camera, **HS_ImageList**, and the list of the files containing the hypercubes, **HS_DataList**) and one output (the input list containing the images returned by the hyperspectral camera added with the RGB images of the surfaces of the hypercubes retrieved from the reflectance spectra).

NOTE: **HS_RGB_Rebuild** exploits the functions developed by Jeff Mather[27] to apply the D65 illuminant and 1931 observer from CIE to the data.

1.4.5. Manually select some reference spectra on the surfaces of the test samples (**White Circles** and **Numbers** in **Figure 2, column A**) by running the isolated measuring points selection code, **PointSel**. Select the measurement points by clicking an interactive window that shows, one by one, the 2D, RGB images of the FOV(s). Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

```
[References] = PointSel(HS_DataList, HS_ImageList);
```

1.4.5.1. There are two inputs (the list containing the images returned by the hyperspectral camera, **HS_ImageList**, and the list of the files containing the hypercubes, **HS_DataList**) and one output (a variable, References, containing the spectra selected as references within the FOV(s)).

1.4.6. If desired, store the position of the references over the surface of the samples into a set of pictures by running the dedicated code, **SaveImPoint**. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

```
SaveImPoint(References, HS_ImageList);
```

1.4.6.1. There are two inputs (the variable containing the reference spectra, **References**, and the list containing the images returned by the hyperspectral camera, **HS_ImageList**) and no outputs (the code saves .png images in the current work folder).

1.4.7. Organize the references into a matrix by running the conversion code, **RefListToMatrix**. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

```
[References_Matrix] = RefListToMatrix(References, HS_ImageList(1).WaveL);
```

1.4.7.1. There are two inputs (the variable containing the reference spectra, **References**, and the list of the wavelengths at which the detector counts the photons during the data acquisition of the spectra, **HS_ImageList(1).WaveL**) and one output (the same reference spectra organized into a matrix, **References_Matrix**).

NOTE: This step is mandatory because the code that evaluates the SAM maps requires the reference spectra to be organized into a matrix. The syntax of the second input, **HS_ImageList(1).WaveL**, is required to recall the variable **WaveL** from the list **HS_ImageList**. The number 1 within brackets refers to the first element of the list named as **HS_ImageList**;

however, since all the hypercubes have the same wavelength range, it can be substituted by each number minor or equal to the total number of listed images.

1.4.8. Extract the SAM maps using the whole spectra by running the standard SAM maps evaluation code, **SAM_Standard**. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

SAM_Standard(HS_ImageList, HS_DataList, References_Matrix);

1.4.8.1. There are three inputs (the list containing the images returned by the hyperspectral camera, **HS_ImageList**; the list of the files containing the hypercubes, **HS_DataList**; and the matrix of the reference spectra, **References_Matrix**) and no output: the code saves the SAM maps as .png images in the current work folder.

1.4.9. Check whether the obtained similarity maps (**Figure 2, columns B–E**) display the details of the brush strokes used to realize the test samples. If this is the case, move to the next step of the protocol; otherwise go back to step 1.3.1 and restart.

2. Adjust the experimental parameters to the painting

2.1. Identify the Region(s) of Interest, ROI(s), of the painting to be studied (red rectangles in **Figure 3A**).

NOTE: It is common that more than one FOV is necessary to cover a single ROI.

2.2. Arrange the acquisition setup and the painting at the distance defined in the previous steps and perform the white calibration employing the white standard reference supplied with the hyperspectral camera.

NOTE: If the user must do an *in situ* acquisition (i.e., with a painting exposed in a museum or at an exhibition), manage only the camera. This is the case of *Quarto Stato*, which is permanently exposed in a dedicated space at the Museo del Novecento in Milan, Italy.

2.3. Acquire the hyperspectral data from at least one FOV within the edge of each ROI(s) (unshaded areas within the red rectangles in **Figure 3A**).

2.4. Download the files returned by the instrument and save them into a dedicated folder.

2.5. Check whether the illumination of the surface of the painting has been uniformly set by looking at the RGB images returned by the hyperspectral camera. If this is the case, move to the next steps, otherwise go back to step 2.2 and restart.

NOTE: **Figure 4** illustrates the importance of this check (see the **Discussion** section for the details).

2.6. Repeat the sub-steps of step 1.4.

2.7. Check whether the data have a spatial resolution high enough to distinguish the brush strokes by observing the RGB pictures of the FOVs (**Figure 3B**) and the SAM maps (**Figure 3C**) related to the reference spectra selected within the FOVs (green circles in **Figure 3B**).

2.8. If the illumination and the spatial resolution have been properly set, complete the collection of the data acquiring the other FOVs necessary to cover the ROI(s); otherwise go back to step 2.2 and restart.

NOTE: When a ROI requires more than one FOV to be covered, ensure a certain degree of superposition between adjacent FOVs to easily stitch the resulting maps[3,15]. The extent of the overlapping depends on the distance between the hyperspectral camera and the sample, on the translation, and the horizontal angle of view of the detector[28]. In the case of the experimental campaign conduced on *Quarto Stato*, the overlap has been set to be at least the 40% of the FOVs.

3. Hypercubes and the reference spectra management

3.1. Perform the I/O of the raw data: organize, read, and manage the hypercubes.

3.1.1. Run the **HS_FileLister** code to store the list of the files containing the hypercubes and the related information into two variables at the disposal of the algorithm (see step 1.4.2 for the practical details).

NOTE: The hyperspectral camera returns hdr (high dynamic range) files that the code manages exploiting a revisited version of the script developed by Jarek Tuszinsky[29].

3.1.2. Run the **HS_Crop_png** code to select the portion of each FOV to be used in the analysis of the data (see step 1.4.3 for the practical details).

3.1.3. Run the **HS_RGB_Rebuild** code to retrieve the RGB images of the FOVs from the reflectance spectra (see step 1.4.4 for the practical details).

3.2. Organize, read (if required), and manage the reference spectra.

NOTE: The reference spectra will play the role of the end members within the SAM method[20,21]. This part of the algorithm is not univocally determined but depends on the selection mode and on the origin of the reference spectra.

3.2.1. Run the **PointSel** code and click within the displayed interactive window to identify the reference spectra as isolated measuring points over the surface of the monitored area(s) (**Figure 5A**) (see step 1.4.5 for the practical details).

3.2.2. Automatically select the reference spectra as a regular reticulum of measuring points superimposed to the surface of the monitored area(s) by running the reticular selection code, **ReticularSel** (**Figure 5B**). Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

[References] = ReticularSel(HS_DataList, HS_ImageList, n_pixel);

3.2.2.1. There are three inputs (the list containing the images returned by the hyperspectral camera, **HS_ImageList**; the list of the files containing the hypercubes, **HS_DataList**; and the spacing of the reticulum, **n_pixel**, expressed in number of pixels) and one output: a variable containing the spectra selected as references within the FOVs, References.

3.2.3. Run the external references importer code, **Spectra_Importer**, to create a variable containing references from datasets and databases independent from the hypercubes acquired on *Quarto Stato*. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

[ExtReferences] = Spectra_Importer (file_extension);

3.2.3.1. There is one input (the extension of the file containing the independent reference spectra, **file_extension**, written between apices) and one output (a variable containing the external references, **ExtReferences**).

NOTE: The external reference importer code has been optimized for importing tmr files but, if necessary, it can be easily modified to deal with any kind of text file.

3.2.4. Run the **RefListToMatrix** code to put the references into a matrix, **References_Matrix** or **ExtReferences_Matrix**, as required by the code that evaluates the SAM maps (see step 1.4.7 for the practical details).

3.2.5. Wait for the **RefListToMatrix** code to equalize both the wavelength range and the spectral resolution (i.e., the number of components) of the hypercubes and the references.

NOTE: The code identifies the wavelength ranges of both the hypercubes and the references. The code compares the wavelength ranges and cuts off the wavelength interval(s) that are not monitored by both the hypercubes and the references. The code identifies the group of hyper-vectors (the hypercubes or the references) constituted by the lower number of components (i.e., characterized by the lower spectral resolution) in the equalized wavelength range. The code reduces the number of components of the longer hyper-vectors (the references or the hypercubes) to that of the shorter ones (the hypercubes or the references). This is done by keeping, for each wavelength of the shorter hyper-vectors, only the values of the longer hyper-vectors that correspond to the nearest wavelength to that of the shorter hyper-vectors.

3.2.5.1. The code automatically performs the equalization. If the references have been selected within the hypercubes, the wavelength range and the spectral resolution do not need to be equalized and they remain unchanged.

3.2.6. If desired, store the position of the references over the surface of the samples into a set of pictures by running the dedicated code (see step 1.4.7 for the practical details).

NOTE: This option is available only if the references have been selected within the hypercubes (steps 3.2.1 and 3.2.2).

4. SAM analysis

4.1. Run the **SAM_Complete** code to evaluate the similarity maps. Type the following command line (semicolon included) in the terminal window and press **Enter** to run the code:

SAM_Complete(HS_ImageList, HS_DataList, References_Matrix);

4.1.1. There are three inputs (the list containing the images returned by the hyperspectral camera, **HS_ImageList**; the list of the files containing the hypercubes, **HS_DataList**; and the reference matrix, **References_Matrix** or **ExtReferences_Matrix**) and no outputs (the code saves the SAM maps as .png files in the current work folder).

NOTE: Other than the three described input variables, the **SAM_complete** code must be fed with few additional parameters to tailor the analysis protocol according to the preferences of the user (see the next steps).

4.2. When required, feed the code with the pre-processing option by typing the number 0 or 1 in the terminal window depending on the desired pre-processing operation and press **Enter** to continue.

4.2.1. Pre-processing option set to 0: the area subtended by each reflectance spectrum is normalized to 1.

4.2.2. Pre-processing option set to 1: the area subtended by each reflectance spectrum is normalized to 1 and then the normalized spectrum is derived one time.

NOTE: Both the hypercubes and the references undergo the same pre-processing option.

4.3. Select the end members to be used for the SAM analysis among the reference matrix by feeding the code with the numbers of the columns that correspond to the desired spectra. When required, enter into the terminal window the sequence of numbers corresponding to the desired columns by typing the numbers within square brackets and separated by a white space. Press **Enter** to continue.

NOTE: The sequence [1 2 3] corresponds to the selection of the first three columns of the reference matrix; an empty vector corresponds to the selection of all the columns of the reference matrix.

4.4. Feed the code with a string containing the first part of the name that will identify the sets of maps to be saved (i.e., the common part of the name of the .png files returned by **SAM_Complete**). When required, insert the string in the terminal window by typing the letters within apices. Press **Enter** to continue.

NOTE: If the user types test, then the name of all the output .png images will start with test.

4.5. When required, feed the code with the method selected to handle the data by typing the number 0, 1, or 2 in the terminal window depending on the desired handle method and press **Enter** to continue.

4.5.1. Set the method to 0 for no manipulation of the data.

4.5.2. Set the method to 1 to require manual selection of the wavelength range(s) of the spectra to be considered before starting the analysis (**Figure 6**).

4.5.3. Set the method to 2 to require the algorithm to order the data on the basis of a specific criterion before the evaluation of the SAM maps (**Figure 7**).

4.6. Wait for the protocol to process the data and to save the SAM maps in the current work folder as .png files.

NOTE: If the handle method has been set to 0 or 2, the user must just wait. If it has been set to 1, the user must select the portion(s) of the spectra to be employed for evaluating the SAM maps by clicking on an interactive window (**Figure 6**).

Representative results:

The proposed protocol offers a set of interesting features for the management and the analysis of HRI data. The I/O (step 3.1) of the raw data is always the first problem that must be solved before applying any analysis method and it can become a critical issue when dealing with large amounts of data. In the present case, the only task regarding the raw data is to store the experimental results into a dedicated folder and select it by browsing the hard disk when running the reading code (step 3.1.1). Thereafter, the cropping and the RGB rebuilding codes allow the refinement of the selection of the data to be analyzed (step 3.1.2) and checks that the experimental conditions have been properly set at the moment of the acquisition of the hypercubes (step 3.1.3, see **Figure 4** and the **Discussion** section for further details).

Once verified that the data-cubes have been correctly acquired, the algorithm offers different possibilities to select the end members for the SAM analysis[20,21] (step 3.2). The first two options (steps 3.2.1 and 3.2.2) retrieve the references among the hypercubes by manually selecting some isolated measuring points (**Figure 5A**) or by automatically sampling the surface of the painting providing a reticular selection of measuring points within one or more FOVs (**Figure 5B**). The analysis based on isolated measuring points is faster than the reticular based one, but it implies a careful and, possibly, informed observation of the FOV(s) to identify the significant spectra; this means good experience dealing with pigments and painted surfaces. The reticular selection makes the algorithm time-consuming and forces the user to observe a lot of output images to retrieve a handful of useful similarity maps. However, the reticular selection provides a complete screening of the hypercubes and, mostly, it can be carried out without experience of the experimental context. In principle, once the sampling distance, **n_pixel**, is decided, the user can neglect the observation of the FOV(s) with a very low probability of losing details.

In addition to the selection of the reference spectra within the hypercubes, the algorithm offers the opportunity to compare the data from the sample under investigation with references belonging to other sources (step 3.2.4). The external reference spectra importer code manages the I/O of references that do not belong to the surface of the painting. The matrix converter code equalizes the wavelength ranges and the spectral resolution of both the hypercubes and the external references (step 3.2.4). This possibility extends the capabilities of the user regarding the characterization of the sample. Indeed, the user can exploit every kind of available resource in terms of reflectance data. The hypercubes can be compared with public databases, with the spectral archives of the user, with new data collected on *ad hoc* prepared samples or even on other objects (paintings, palettes, hues, or whatever) belonging to the author or to other artists. Moreover, the external references can be obtained exploiting any kind of reflectance techniques so much so that the references shown in **Figure 6** and **Figure 7** have been acquired by a portable FORS miniature spectrometer (**Table of Materials**) and not by the camera used for the HRI data.

Beyond the data management, the algorithm offers an original approach to the data analysis too. It allows manipulation of the spectra before evaluating the SAM maps (steps 4.1–4.5). This possibility finds its rationale in the choice of the SAM method to investigate the distributions of the pigments. In fact, SAM considers the reflectance spectra as they would be vectors in a multi-dimensional space (i.e., hyper-vectors with a number of components equal to that of the acquisition channels). Therefore, if the principal aim of the analysis is to compare different but similar references to distinguish which one best matches the pigments used by the artist, then the almost identical components of the reference spectra (i.e., the wavelengths that correspond to almost identical values in the hyper-vectors) should not be particularly useful and the algorithm allows to exclude these components from the analysis.

The protocol supports two options for manipulating the data (step 4.5): the user can define the wavelength portion(s) of the reflectance data to be analyzed manually (**Figure 6**) or automatically (**Figure 7**). The manual selection is straightforward. The pre-processed reference spectra or their first derivatives, depending on the selected pre-processing option (step 4.2), appear on an interactive window, **Figure 6A**, and the user selects one or more wavelength interval(s), **Figure 6B**, by clicking on the graph surface. The automatic selection is based on the mathematical criterion of the maximum variance applied to the pre-processed reference spectra or their first derivatives, depending on the selected pre-processing option (step 4.2). The algorithm computes the variance (normalized and displayed as a **Dashed Line** in **Figure 7A**) within the selected references and order all the spectra (both the references and the hypercubes) accordingly to this criterion (the **Dashed Line** in **Figure 7B** represents the normalized and ordered variance). In other words, if the maximum variance corresponds to the nth wavelength, the content of the nth component of each pre-processed spectrum (references and hypercubes) will be moved to the first position of a re-arranged hyper-vector and so on (the colored portions of the background in **Figure 7A** and **Figure 7B** graphically explain the re-arrangement of the data). Practically, the components of the pre-processed spectra are ordered similar to principal component analysis[30].

Once the spectra have been manipulated, the algorithm evaluates the SAM maps. Following the manual manipulation (**Figure 6**), the protocol returns three sets of maps: two corresponding to the groups of selected and rejected wavelengths and one obtained employing the whole spectra. Otherwise, following the automatic manipulation (**Figure 7**), the algorithm applies a floating threshold to the variance values and evaluates the SAM maps at the increasing threshold for both the re-arranged hyper-vector components corresponding to the over threshold (i.e., automatically selected) and to the under threshold (i.e., automatically rejected) values of the variance. These sets of maps, together with that obtained from the whole spectra (always returned by the algorithm), result in a total of (2N + 1) sets of maps where N is the number of values assumed by the threshold. The sets of similarity maps obtained at the increasing of the threshold (**Figure 8**) illustrate that data manipulation does not alter the content but rather provides new insights into the details of the mapped area(s) and, consequently, can help to distinguish similarities and differences between the samples and the references.

Figure and table legneds:

Figure 1: ***Quarto Stato.*** A picture of the painting, 1899–1901, 293 x 545 cm, oil on canvas, Giuseppe Pellizza da Volpedo, Museo del Novecento, Milan, Italy.

Figure 2: Definition of the experimental conditions. (**A**) The *ad hoc* prepared test samples; the **White Circles** and **Numbers** identify the measuring points corresponding to the spectra selected as references. (**B**) The SAM maps evaluated with respect to reference spectrum number 1, (**C**) number 2, (**D**) number 3, and (**E**) number 4. The **Gray Color Bar** indicates the range of values of the spectral angles.

Figure 3: The application of the defined experimental conditions to *Quarto Stato*. (**A**) The ROIs selected for the experimental campaign (**Red Rectangles**); in each rectangle a FOV of those necessary to cover the ROI has been highlighted (**Unshaded Areas**). (**B**) The RGB pictures of the four **Unshaded Areas** of panel A. (**C**) The SAM maps evaluated with respect to a reference spectrum selected within each FOV (**Green Circles**). The **Gray Color Bar** indicates the range of values of the spectral angles.

Figure 4: Proper versus improper illumination of the surface of the sample. (**A**) A portion of the FOV where a small fraction of the painted surface (**Red Circle**) is affected by altered reflectance properties due to improper illumination. (**B**) The same small fraction of the painting (**Blue Circle**) as it results when the FOV is properly illuminated. (**C**) The reflectance spectra of the measuring point at the center of the circles when the FOV is improperly and properly illuminated (**Red** and **Blue Line** respectively). (**D**) The SAM map of the FOV obtained using the spectrum of the improper illuminated measuring point as reference. (**E**) The SAM map of the FOV obtained using the spectrum of the proper illuminated measuring point as reference. The **Gray Color Bar** refers to (**D**) and (**E**) and indicates the range of values of the spectral angles obtained comparing the first derivatives of the spectra of the hypercube of the selected FOV and the first derivative of the spectrum of the measuring point at the center of the colored circles in (**A**) and (**B**).

Figure 5: References selection within the hypercubes. (**A**) The isolated measuring points selection mode; the **Green Circles** indicate the location of the reference spectra manually selected on the FOV shown. (**B**) The reticular selection mode; the **Green Circles** indicate the location of the reference spectra selected by applying a reticulum with the sampling interval

(**n_pixel**) set to five pixels to the FOV shown. The image reported in both (**A**) and (**B**) is the grayscale conversion of the RGB image of the FOV retrieved from the reflectance spectra applying the D65 illuminant and 1931 observer from CIE standards to the hypercube; the **Gray Color Bar** refers to the normalized intensity of this image.

Figure 6: The manual data manipulation mode. (**A**) The aspect of the interactive window that allows the user to divide the reference spectra into the selected and rejected fractions of wavelengths. (**B**) The same references of (**A**) where the portions of data selected for evaluating the SAM maps have been highlighted by a pink background. (**A**) and (**B**) display the pre-processed spectra of the references.

Figure 7: The automatic data manipulation mode. (**A**) The first derivatives of the four normalized references reported in **Figure 6** (**Coloured Lines**) and their normalized maximum variance (**Black Dashed Line**). (**B**) The same derivatives of (**A**) sorted following the criterion of the maximum variance; the sorted values of the normalized maximum variance have been reported too (**Black Dashed Line**). Some portions of the background have been colored with different hues in the attempt to visually illustrate the re-arrangement of the hyper-vectors.

Figure 8: The SAM maps obtained by the automatic data manipulation mode. (**A–C**) The sorted values of the normalized maximum variance evaluated within the first derivatives of the reference spectra of **Figure 7**; the **Green** and **Red Sections** of the curve indicate, respectively, the selected (over threshold values) and rejected fraction of the data (under threshold values). The panels show what happens at the increasing of the threshold (**Black Dotted Segment**); each panel reports the SAM maps for both the groups of values obtained for the four derivatives of the spectra of **Figure 7**; the **Green Edged Maps** refer to the over threshold fractions while the **Red Edged Maps** refer to the under threshold ones. The **Gray Color Bars** indicate the range of values of the spectral angles. In this example, the step that determines the increase of the threshold is equal to 0.5% of the normalized maximum variance; the threshold values reported in (**C**) is 0.09 and it is the last considered threshold value because a further increase would reduce the number of selected components of the hyper-vectors below the arbitrarily fixed lower limit of 20 values, i.e., 10% of the total number of the acquisition channels of the hyperspectral camera.

Discussion:

Hyperspectral reflectance imaging datasets are large reservoirs of information; therefore, the development of robust and, possibly, automated protocols to analyze the data is a key turn to exploit their potential[15,17]. The proposed algorithm answers this need in the field of cultural heritage with particular attention to the characterization of the pigments of paintings. Based on SAM[20,21], the algorithm supports the user during the whole analysis process from the setting of the experimental conditions to the evaluation of the distribution of pigments. Though the algorithm still does not have a complete graphical interface and that it does not provide a tool for viewing the results (for this purpose an open-source software has been used[31] and it is recommended, see **Table of Materials**), the set of possibilities implemented to modulate the approach to the data analysis extensively balances these drawbacks.

The protocol sets the acquisition system according to the characteristics of both the sample and the detector. On one hand, the Divisionist technique employed by Pellizza Da Volpedo to create *Quarto Stato* requires that the hypercubes distinguish between small brush strokes of different pigments placed side by side. On the other hand, the hyperspectral camera has a focus range between 150 mm and infinite with a manual adjustment system and a 1 m distance to the target that detects an area of 0.55 x 0.55 m with a spatial resolution of 1.07 mm[26]. The application of the algorithm to the few hypercubes acquired on the test samples (**Figure 2**) helps establish a suitable working distance for the data acquisition. The measurements on the test samples were a defined 30 cm, corresponding to a resolution of 0.31 mm at the target. This working distance was also successfully adopted during the experimental campaign conducted on *Quarto Stato* (**Figure 3**). Once the working distance was defined, the illumination of the surface of the sample remains a critical issue[3,15]. When a portion of a FOV shows uneven (**Red Circle** in **Figure 4A**) instead of uniform illumination (**Blue Circle** in **Figure 4B**), the reflectance properties change dramatically (**Figure 4C**) and the whole procedure is compromised (**Figure 4D** vs. **Figure 4E**). The protocol prevents uneven illumination (and more in general against artefacts in the monitored areas) during the acquisition of the data (by returning RGB in step 1.4.1 and SAM maps in step 1.4.9 that can be checked by the user) and *a posteriori* by excluding the compromised portions of the FOVs from the analysis by means of the cropping code (steps 1.4.2 and 3.1.2).

The protocol allows the user to select the references (i.e., the end members used for the evaluation of the SAM maps) with the maximum freedom. On one hand, the EMs can be chosen within the edges of the hypercubes by two manners: isolated measuring points selection (**Figure 5A** in steps 1.4.5 and 3.1.2) or reticular measuring points selection (**Figure 5B** in step 3.1.3). The first can be defined as informed selection because it requires some expertise in the user to manually identify the significant measuring points. The latter can be defined as blind selection because the reticular sampling of the FOVs requires only the value of the sampling interval to be performed. On the other hand, the EMs can be retrieved from outside the painting under investigation (step 3.1.4). During the experimental campaign conducted on *Quarto Stato*, a portable miniature FORS spectrometer (**Table of Materials**) was used to collect spectra from draft samples belonging to the artist and currently kept in the Studio Museum located in Volpedo (Pellizza da Volpedo Studio Museo, Volpedo (AL), Italy). These reflectance data have been used for the evaluation of the SAM maps and some of them are reported in **Figure 6** and **Figure 7**. Since it limits the importance of the absolute intensity and of the baseline of the spectra, the pre-processing is mandatory for both the hypercubes and the EMs, especially if they have been obtained from slightly different setups or operative conditions[32].

The last main feature of the protocol is the chance to manipulate the experimental data. For manipulation, it is intended that the identification of the most significant components of the EMs (i.e., of those portions of the spectra) should help characterize the materials used by the artist. This task can be accomplished manually (**Figure 6**) or automatically (**Figure 7**). In the first case, the algorithm takes advantage of the expertise of the performer while, in the second case, it is a statistical criterion that determines the components that, time by time, will be used to evaluate the SAM maps. In both cases, the manipulation increases the number of the resulting similarity maps and consequently extends the capability to disclose the information carried by the hypercubes. In particular, the criterion-based selection generates a great number of insights of the painted surface (**Figure 8**).

Taken individually, the enumerated features could appear as mere technical benefits, but together they imply at least two main key points. The algorithm can be successfully applied by any kind of user and it can significantly broaden the scenario of the analysis. In fact, the main steps of the protocol (i.e., the selection of the references and the manipulation of the data) can be performed automatically, disregarding the skills and the experience of the user. With the possibility to drive the analysis with spectra from outside the hypercubes, all the reflectance data in the disposal of the researchers can be exploited for the characterization of the sample under investigation.

In summary, the protocol can be an extremely flexible tool. With some improvements regarding the graphical interface and the number of supported analysis methods, it can be a step beyond the state of the art regarding the handling and the analysis of data obtained from painted surfaces by means of hyperspectral reflectance imaging.

Acknowledgments:

This research was funded by Regione Lombardia in the framework of the Project MOBARTECH: una piattaforma mobile tecnologica, interattiva e partecipata per lo studio, la conservazione e la valorizzazione di beni storico-artistici - Call Accordi per la Ricerca e l'Innovazione.

The authors are grateful to the staff at Museo del Novecento for the support during the *in situ* experimental sessions and to the Associazione Pellizza da Volpedo for the access to Studio Museo.

Disclosures:

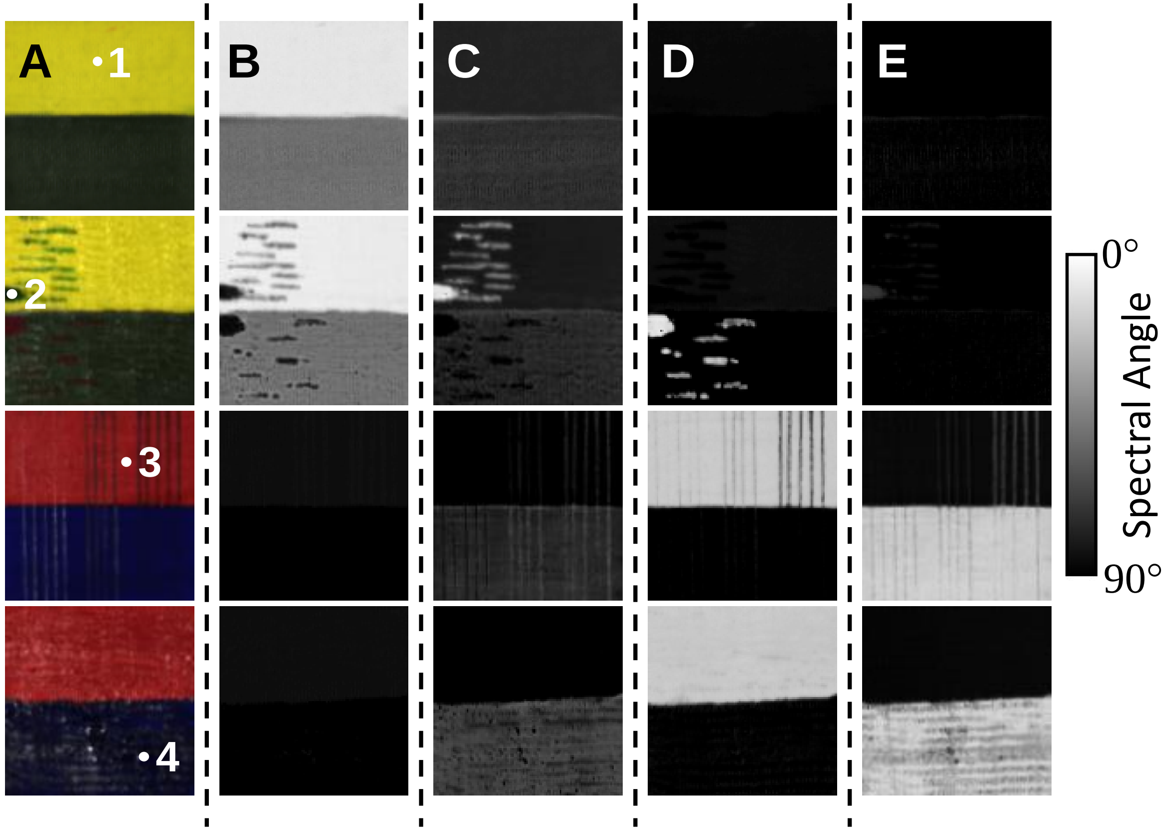
The authors have nothing to disclose.

References:

1. Picollo, M. et al. Fiber Optics Reflectance Spectroscopy: a non-destructive technique for the analysis of works of art. in *Optical Sensors and Microsystems*. Springer. 259–265 (2002).
2. Bacci, M. et al. Non-destructive spectroscopic investigations on paintings using optical fibers. *MRS Online Proceedings Library Archive*. **267** (1992).
3. Liang, H. Advances in multispectral and hyperspectral imaging for archaeology and art conservation. *Applied Physics A*. **106**, 309–323 (2012).
4. Bullock, L. Reflectance spectrophotometry for measurement of colour change. *National Gallery Technical Bulletin*. **2**, 49–55 (1978).
5. Saunders, D. Colour change measurement by digital image processing. *National Gallery Technical Bulletin*. 66–77 (1988).
6. Appolonia, L. et al. Combined use of FORS, XRF and Raman spectroscopy in the study of mural paintings in the Aosta Valley (Italy). *Analytical and Bioanalytical Chemistry*. **395**, 2005–2013 (2009).
7. Pouyet, E. K. et al. New insights into Pablo Picasso's La Miséreuse accroupie using X-ray fluorescence imaging and reflectance spectroscopies combined with micro-analyses of samples. *SN Applied Sciences*. **2**, 1–6 (2020).

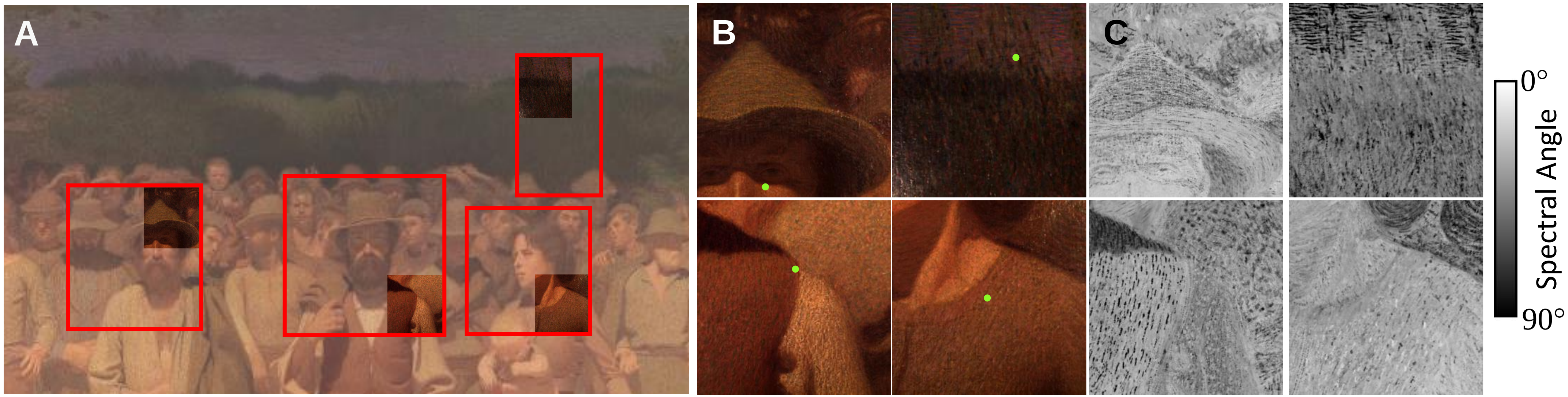
8. Garofano, I., Perez-Rodriguez, J. L., Robador, M. D., Duran A. An innovative combination of non-invasive UV–Visible-FORS, XRD and XRF techniques to study Roman wall paintings from Seville, Spain. *Journal of Cultural Heritage*. **22**, 1028–1039 (2016).
9. Dupuis, G., Elias, M., Simonot, L. Pigment identification by fiber-optics diffuse reflectance spectroscopy. *Applied Spectroscopy*. **56**, 1329–1336 (2002).
10. Bacci, M., Picollo, M. Non-destructive spectroscopic detection of cobalt (II) in paintings and glass. *Studies in Conservation*. **41**, 136–144 (1996).
11. Cosentino, A. FORS spectral database of historical pigments in different binders. *E-Conservation Journal*. 54–65 (2014).
12. Leona, M., Winter, J. Fiber optics reflectance spectroscopy: a unique tool for the investigation of Japanese paintings. *Studies in Conservation*. **46**, 153–162 (2001).
13. Cheilakou, E., Troullinos, M., Koui, M. Identification of pigments on Byzantine wall paintings from Crete (14th century AD) using non-invasive Fiber Optics Diffuse Reflectance Spectroscopy (FORS). *Journal of Archaeological Science*. **41**, 541–555 (2014).
14. Kubik, M. Hyperspectral imaging: a new technique for the non-invasive study of artworks. *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage*. **2**, 199–259 (2007).
15. Fischer, C., Kakoulli, I. Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. *Studies in Conservation*. **51**, 3–16(2006).
16. Daniel, F. et al. Hyperspectral imaging applied to the analysis of Goya paintings in the Museum of Zaragoza (Spain). *Microchemical Journal*. **126**, 113–120 (2016).
17. Baronti, S., Casini, A., Lotti, F., Porcinai, S. Principal component analysis of visible and near-infrared multispectral images of works of art. *Chemometrics and Intelligent Laboratory Systems*. **39**, 103–114 (1997).
18. Mansfield, J. R. et al. Near infrared spectroscopic reflectance imaging: supervised vs. unsupervised analysis using an art conservation application. *Vibrational Spectroscopy*. **19**, 33–45 (1999).
19. Clodius, W. B. Multispectral and Hyperspectral Image Processing, Part 1: Initial Processing. *Encyclopedia of Optical Engineering: Las-Pho, pages 1025-2048*. **2**, 1390 (2003).
20. Kruse, F. A. et al. The spectral image processing system (SIPS)-interactive visualization and analysis of imaging spectrometer data. in *AIP Conference Proceedings* (1993).
21. Yang, C., Everitt, J. H., Bradford, J. M. Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). *Transactions of the ASABE*. **51**, 729–737 (2008).
22. Delaney, J. K. D. et al. 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. *Heritage Science*. **6**, 1–12 (2018).
23. Richards, J. A. Remote sensing digital image analysis. Springer. **3** (1999).
24. Mobaraki, N., Amigo, J. M. HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems*. **172**, 174–187 (2018).
25. Zhu, C. Y. et al. Optimization of a hyperspectral imaging system for rapid detection of microplastics down to 100 µm. *MethodsX*. **8**, 101175 (2021).
26. Behmann, J. et al. Specim IQ: evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. *Sensors*. **18**, 441 (2018).
27. Mather, J. Spectral and XYZ Color Functions (https://www.mathworks.com/matlabcentral/fileexchange/7021-spectral-and-xyz-color-functions). *MATLAB Central File Exchange* (2021).

28. Chen, C.-Y., Klette, R. Image stitching—Comparisons and new techniques. in *International Conference on Computer Analysis of Images and Patterns* (1999).
29. Tuszynski, J. read_envihdr (https://www.mathworks.com/matlabcentral/fileexchange/ 38500-read_envihdr). *MATLAB Central File Exchange* (2021).
30. Jolliffe, I. T., Cadima, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. **374**, 20150202 (2016).
31. Schneider, C. A., Rasband, W. S., Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*. **9**, 671–675 (2012).
32. Rinnan, Å., Van Den Berg, F., Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry*. **28**, 1201–1222 (2009).


A
•1
•2
•3
•4
B
C
D
E
0°
Spectral Angle
90°

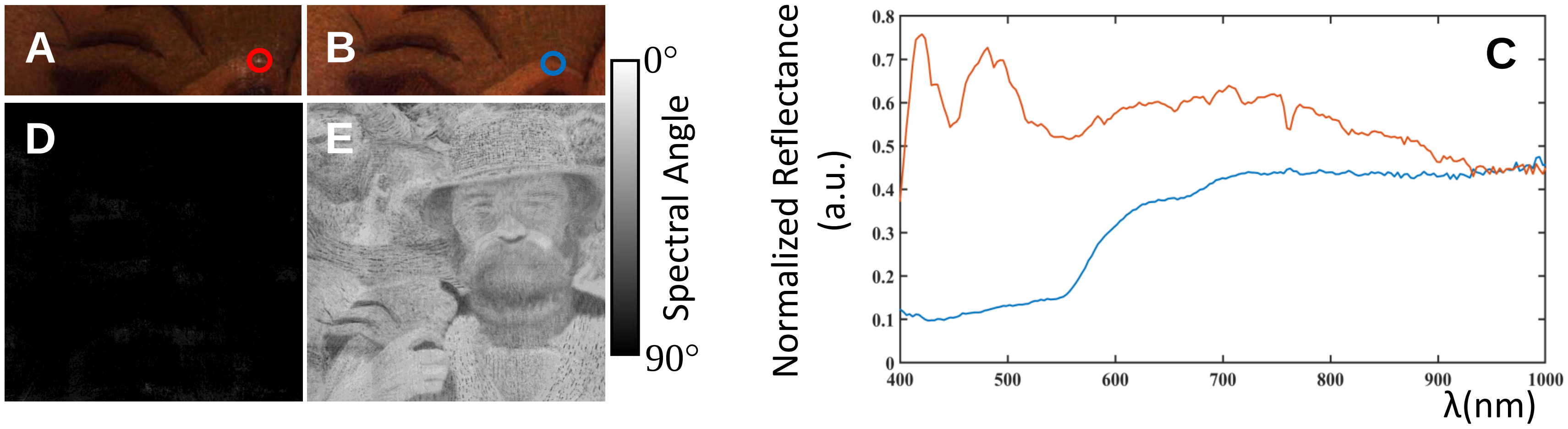
Figure

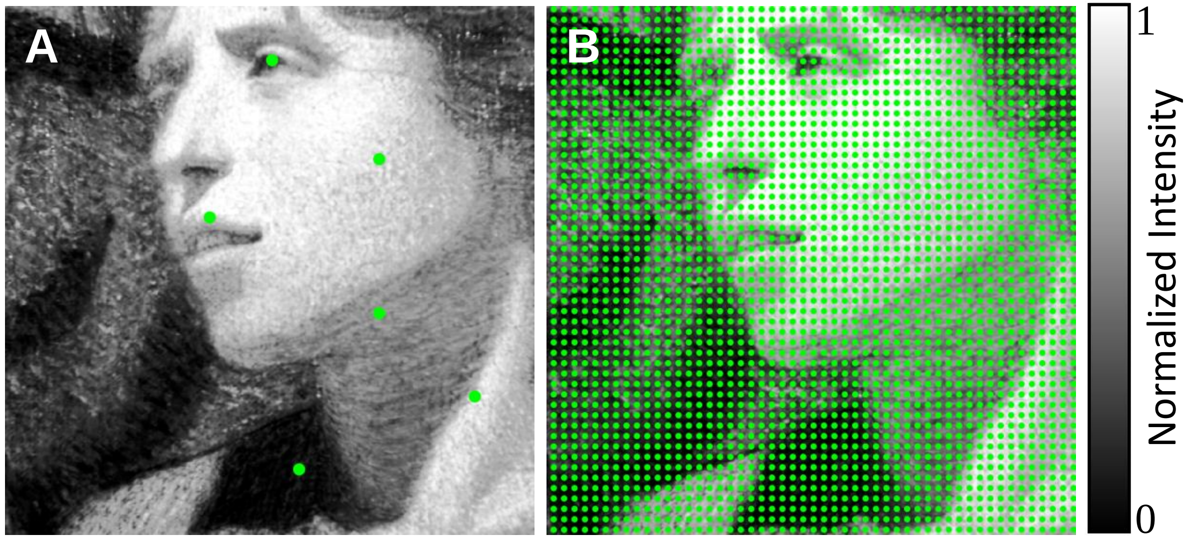
Click here to access/download;Figure;Fig03_05.pdf



Figure

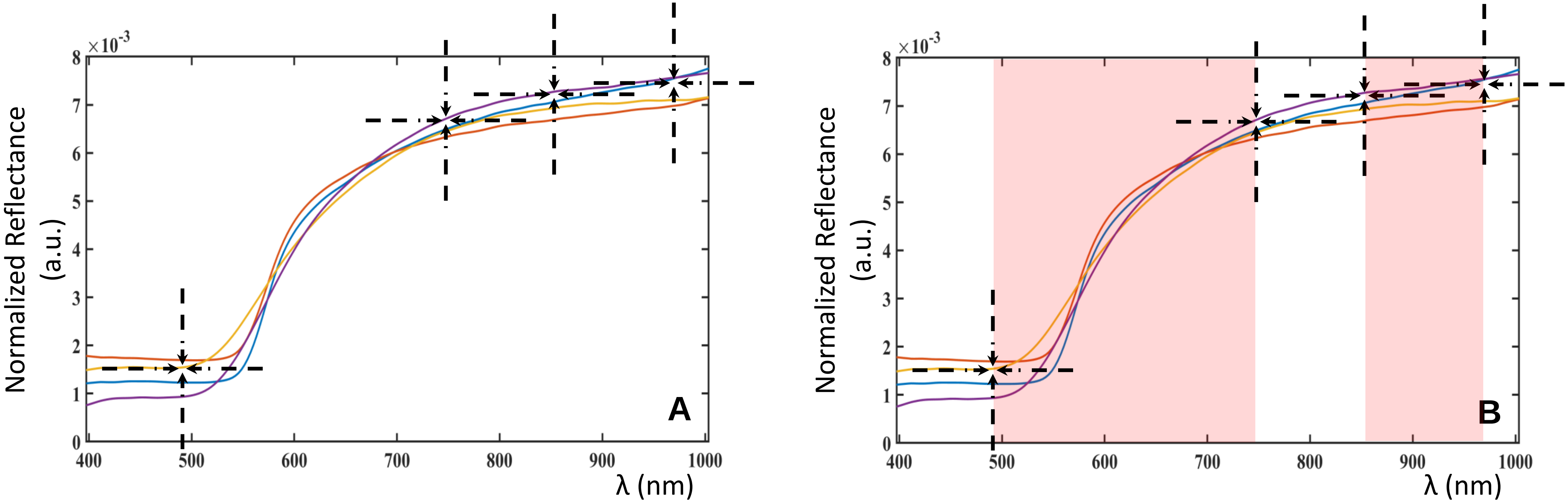
Click here to access/download;Figure;Fig04_05.pdf



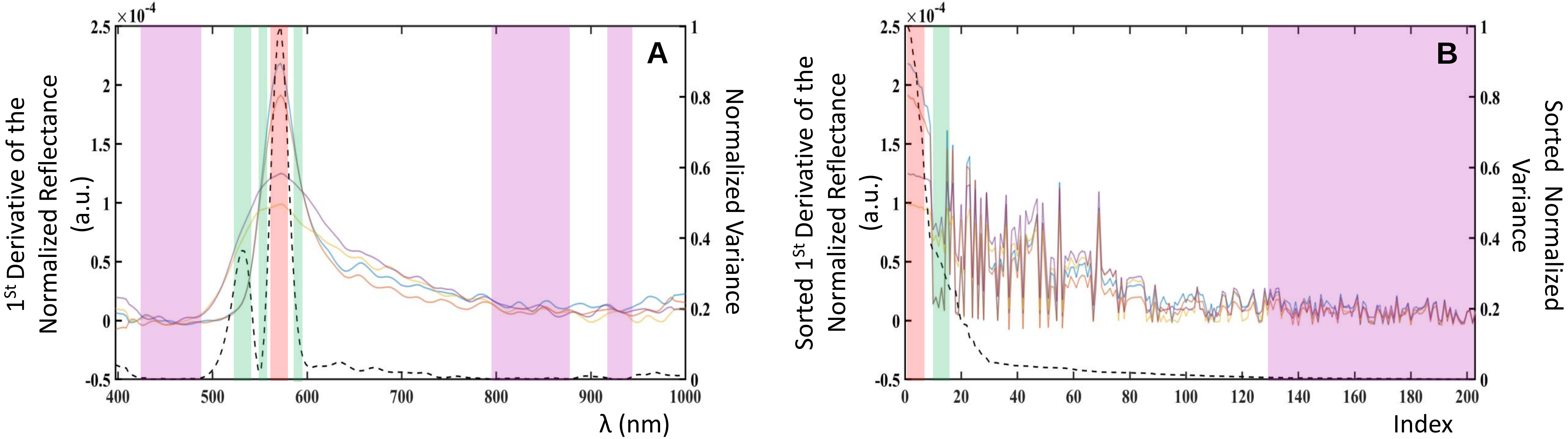

A
B
1
Normalized Intensity
0

Figure

Click here to access/download;Figure;Fig06_05.pdf



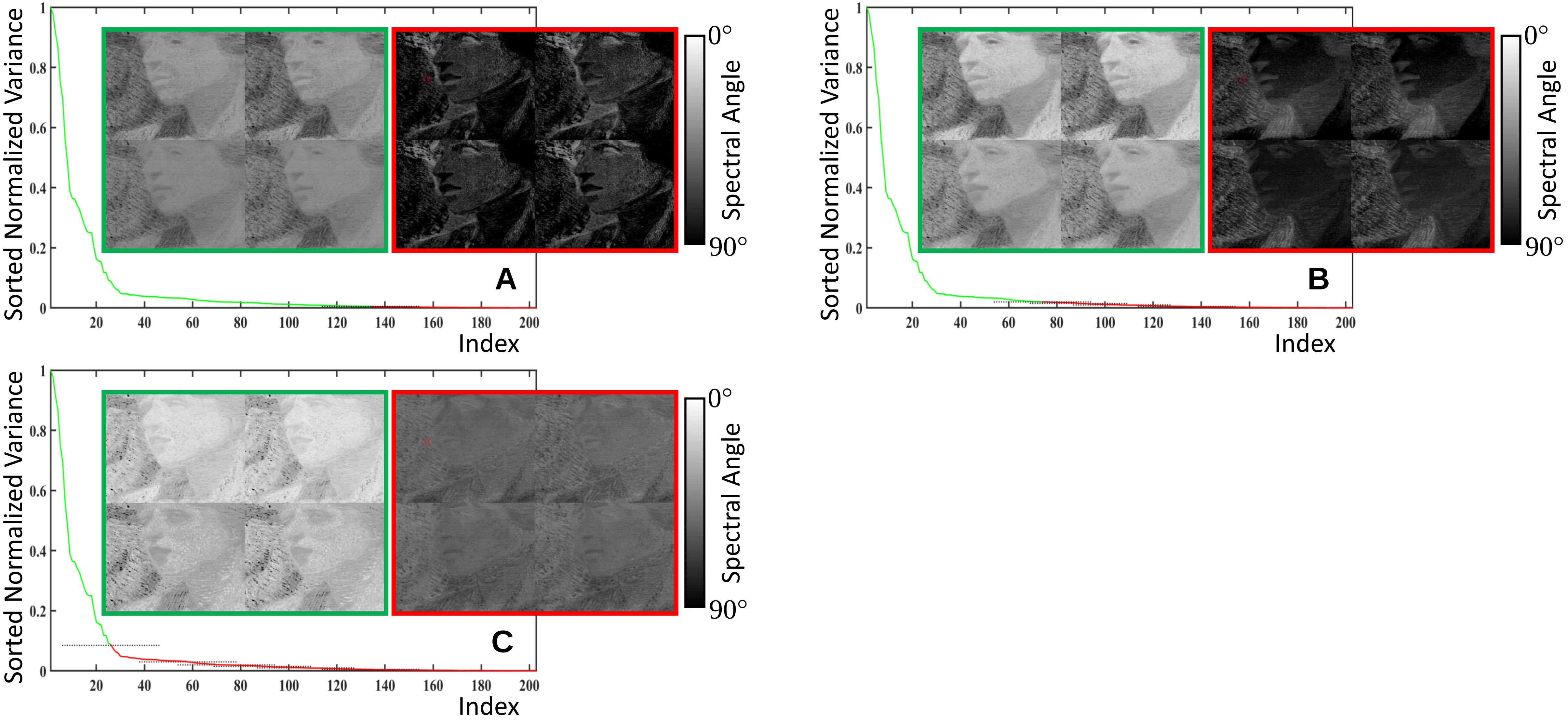
Figure



A
B
×10^-4
2.5
2
1.5
1
0.5
0
-0.5
400
500
600
700
800
900
1000
λ (nm)
1st Derivative of the Normalized Reflectance (a.u.)
Normalized Variance
1
0.8
0.6
0.4
0.2
0
Sorted 1st Derivative of the Normalized Reflectance (a.u.)
Sorted Normalized Variance
20
40
60
80
100
120
140
160
180
200
Index

Figure

Click here to access/download;Figure;Fig08_05.pdf



Name of Material/ Equipment	Company	Catalog Number
ImageJ/Fiji	Specim (Oulo, Finlad)	N/A
MATLAB 2019b	StellarNet Inc (Tampa, Florida, USA)	N/A
Specim IQ Hyperspectral Camera	National Institutes of Health (Bethesda, Maryland, USA)	N/A
StellarNet BLUE-wave Miniature Spectrometer	MathWorks (Natick, Massachusset, USA)	N/A

Comments/Description
Portable reflectance hyperspectral camera used to acquire the hypercubes
Portable reflectance spectrometer used to acquire independent reflectance spectra
Open source Java image processing program
Program Language and numerical computing environment

Dear members of the JoVE editorial team.

Dear Referees

Thank you for the attention you reserved to our manuscript and for the possibility to submit an improved version of our work.

We appreciated the comments of the editorial board and of the referees; they really helped us to make our work better; we re-viewed the manuscript to satisfy their requests and to include their suggestions.

We rationalized the description of the instructions every time the user must run the codes that constitute the computational part of the protocol; the input and output variables have been listed and shortly described. We limited the use of the notes to the bare minimum. We included all the missed details highlighted by the referees; we improved the text where requested and we tried to clarify the ambiguous points; we also improved the literature including new references.

We hope to be notified soon about our work.

Thank you in advance for you time and kindness.

Best regards.

Michele Caccia

New Editorial comments:

1. Please employ professional copyediting services as the language in the manuscript is not publication grade. Additionally, please break up introduction and results into paragraphs.

We broke up introduction and results into paragraphs.

We apologize for our English; we are not mother tongue but in our opinion the language of the text is not so bad to justify the rejection of the manuscript. If you refer to some expressions, please, give us examples so that we can understand what you mean. Anyway, we tried to improve our work.

2. Please revise the summary to be 50 words or fewer

We reduced the summary below the 50 words.

3. Please superscript all reference numberings without brackets.

We adjusted the format of the references according to your request.

4. Please specify how all of the codes are performed:

[HS_DataList HS_ImageList] = HS_FileLister

Are these terminal line commands? Where are they entered?

Yes, they are Matlab terminal line commands and the user must insert them in the Matlab command window keeping the names of the codes as they are written in the protocol. As pointed out in the previous review we chose suitable names for the variables, but the user can change them as he prefers. In this example HS_FileLister is the name of the code and cannot be changed. HS_DataList and HS_ImageList can be substituted by whatever name but they must be included between square brackets and separated by a white space.

We added the instructions to manage the input and output variable at the end of the introduction.

5. Can the codes be provided?

Yes, we can provide the codes (and a set of test data if it will be necessary) but now we do not see a reason to do that. No one of the four referees asked for them; moreover, you must consider that to keep save the intellectual property of the developer is our duty. We suggest rediscussing this issue after that the work has been revised by the referees.

Editorial comments:

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.

We reviewed the text and we think that it is now improved in these aspects

2. JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names before an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials. For example: Specim IQ, MATLAB, etc.

We deleted all the symbols, names and words that could remind commercial entities.

3. Please add more details to your protocol steps. Please ensure you answer the “how” question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action.

After the review, we think that the description of the steps of the protocol is complete and clear

4. Line 99: Please provide more details regarding the preliminary inspection. What are the parameters inspected? What kind of quantification/ qualifications are required?

We reviewed the text and we clarified that we refer to a semi-qualitative evaluation of the size of the brushes of paint used by the Master to realize the painting

5. Line 110: Please elaborate on how to mimic the peculiar characteristics. What characteristics are classified as peculiar?

The painter's technique has been reproduced by an art expert that realized some samples test with stripes and spots of paint of the same size as those used by Pellizza Da Volpedo in *Quarto Stato*. The aim of this operation was to calibrate the most important experimental parameter: the distance between the painted surface and the detector.

6. Line 124: Please mention how is the portion of the hypercube defined. Is it defined as a volume or area or any other characteristic?

Hypercubes are 3D objects therefore when you select a portion of a hypercube you select a 3D object; nevertheless, the selection is done looking at the 2D RGB image of the surface imaged by the hypercube. We specified this point in the steps of the protocol

7. Line 181: Please specify if there is a minimum number of FOVs required.

No there isn't. The number of FOVs depends on the set-up decided by the user; mainly on the size of the portion(s) of surface to be studied and on the desired resolution (i.e. on the distance between sample and detector).

8. In the JoVE Protocol format, "Notes" should be concise and used sparingly. They should only be used to provide extraneous details, optional steps, or recommendations that are not critical to a step. Any text that provides details about how to perform a particular step should either be included in the step itself or added as a sub-step. Please consider moving some of the notes about the protocol to the discussion section.

We reduced the number of notes at the minimum and, when we were forced to use notes, we were as short as possible. The surviving notes contain reminders to secondary details related to the steps of the algorithm, some short clarification or credits when we exploited tools by other authors.

We followed your suggestion and the description of the input and output parameters related to each code have been transformed in sub-step (the bulleted lists) under the command that must be typed in the Matlab command window when running the code

9. Please highlight up to 3 pages of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol. Remember that non-highlighted Protocol steps will remain in the manuscript, and therefore will still be available to the reader.

We performed the selection.

10. As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3-6 paragraphs with citations:

a) Critical steps within the protocol

b) Any modifications and troubleshooting of the technique

c) Any limitations of the technique

d) The significance with respect to existing methods

e) Any future applications of the technique

We reviewed the text and we think that it now matches your request

11. Please ensure that the Table of Materials contain all the essential supplies, reagents, and equipment. The table should include the name, company, and catalog number of all relevant materials in separate columns in an xls/xlsx file.

We adapted the Table of Material to the required format.

Reviewer #1:

Manuscript Summary:

Creation of a semi-automatic code to create classification maps of hyperspectral data.

Major Concerns:

This paper should have been edited before submission as there are major grammatical errors making reading / understanding difficult.

Thank you for your comment. We reviewed the text and we think that it is now easy to be read

The reason for the creation of test samples are unclear. If the goal is to identify pigments in a painting - how can samples be created without knowledge of the pigments?

In shed of your comment we realized that we have not been clear in the description of this step of the protocol. You will be right if the idea behind the realization of the ad-hoc samples would concern the characterization of the pigments themselves, but here the attention is focused on the size of the brush strokes applied by Pellizza da Volpedo on the canvas. Divisionist painters obtain the desired chromatic effect (i.e. the colour appreciated by an observer looking at the painting by a certain distance) juxtaposing dots or stripes of different hues according to the painter's interpretation of the colour mixing theory. If we want to uncover the materials used by Pellizza, we must be sure to be able to discern between the different thin stripes adopted by the artist. This means that the spatial resolution of the data cubes must be high enough to distinguish the stripes independently from the employed pigments. Therefore, we asked an art expert to create some test samples that mimic the size of the brush strokes used by Pellizza and, since the spatial resolution of the hypercubes is determined by the distance between the Specim IQ camera and the canvas, we used the test samples to adjust the canvas camera mutual positions. We set up the experimental conditions in the lab exploiting the test samples to optimize the timing of the *in situ* experimental campaign. We reviewed the text to better explain this point.

Additionally, why use ROI's and not the complete datacube? With matrix math in Matlab this is not computationally problematic even with large datasets.

Here we probably have a misunderstanding. In the manuscript a ROI does not define a portion of a data cube returned by the IQ camera but rather a portion of the painting selected by the historians for understanding the main features of the surface under investigation; the area covered by a single data cube is referred as a FOV and we stated that we need more FOVs (i.e. more data cubes) to cover each of the ROIs reported in figure 3; for instance we used 9 cubes to cover the ROI around the face of the woman in the front right part of the painting.

It is true that we provide a code (HS_Crop_png, point 1.4.3 and 3.1.2 of the protocol) that allows the user to discard portion of the data cube, but we also explain that it is intended to be used to prevent the analysis from artefacts: in particular we refer to uneven illumination of portions of the FOVs (situation illustrated by figure 4 and, indeed, encountered during the study of Quarto Stato) or to the presence within the FOVs of elements unrelated to the painted surface such as the frame surrounding the painting; this is usually a concrete possibility (especially if the painting is small or if the user is interested to areas at the edges of the painted surface) but, we did not suffer this situation because the ROIs selected for the analysis of Quarto Stato are fully within the painted surface (see figure 3).

There is definitely a need in this field for automated algorithms, but it does not seem this paper addresses this well.

We respect the opinion of the referee, but we disagree with this statement. The protocol is fully independent from the case study, it manages a huge amount of data, it helps to avoid the effects of some of the principal artefacts dealing with reflectance data cubes and it can be tailored on the needs and on the degree of knowledge about works of art of the user just by feeding a code with few parameters. In addition, it defines a new idea for what concerns the statistical analysis of the reflectance data; it is true that an advanced GUI must be designed but, for what concerns the algorithm, the offered features and their importance are quite evident. Anyway we tried to improve the text to further highlight the aspects stressed be the referee.

Reviewer #2:

Manuscript Summary:

The manuscript synthesizes an application protocol using a commercial device (Specim IQ Hyperspectral Camera). The protocol is very interesting and requires some modifications to improve it.

Major Concerns:

One of the main problems with these types of cameras is the repeatability of the measurement in order to obtain the same result.

Specim IQ hyperspectral camera is a snapshot camera without an illumination system. For a repeatable measurement some information is not clear: Lighting system, working distance, pixel size obtained with the working distance, geometry of the light source in order to avoid shadow areas. (I suggest a schematic image

From the tests carried out in the laboratory, with samples made to simulate the Divisionist technique used by the author of the painting, a setting was chosen that uses two lamps placed symmetrically to the right and left of the camera and which illuminate the entire area of interest under examination (ROI) in a homogeneous way. The camera was moved by a tripod with an extendable central column with a rack and equipped with wheels on rails in order to keep it at the same distance from the painting. The lamps are two 500W halogen lamps that can be adjusted by means of a dimmer

The Specim IQ camera has been fully characterized by Behmann and co-workers. the instrument guarantees a spatial resolution of 1.07 mm/pixel with a working distance of 1 m; considering that the cubes returned by the camera are constituted by 512x512x204 voxels the size of each of the 204 matrices of the reflectance cube is 55x55 cm; this correspond to a numerical aperture of about 0.26 for the camera. In the case of Quarto Stato each data cube covers an area of 16x16 cm, it has been obtained with the camera located about 30 cm from the painting surface and it is characterized by a resolution of 0.31 mm/pixel. We included these details in the text (see the discussion section)

Did you use a white reference for the white calibration? if yes enter more information.

Yes, we used the white reference supplied with the Specim IQ camera, the white reference panel material has 100% reflectivity in the wavelength range of the hyperspectral camera as declared by the manufacturer. We inserted this detail in the protocol where appropriate-

Minor Concerns:

Regarding data processing, the device output is a raw image (with black and white reference image for calibration in matlab) or directly a normalized image? Add this information

Thank you for this observation. The Specim IQ camera returns both the raw and the B/W calibrated images. We used the calibrated hypercubes and we included this detail in the text (see the note below step 1.3.2).

Probably it is my carelessness but I not have found the number of wavelengths acquired by the Specim IQ Hyperspectral Camera

Thank you again; we apologize for having missed these important details (spectral and spatial resolution/range); they have been added in the text in the note below step 1.3.3. Luckily, even in the un-reviewed version of the manuscript, the interested reader could have found all the technical characteristics of the camera in the work by Behmann. As reported by Behmann and co-workers, the Specim IQ camera uses 204 channels to monitor the reflectance signal in the wavelength range between 400 and 1000 nm with a spectral resolution of 7 nm at FWHM.

Reviewer #3:

Manuscript Summary:

General comments

This manuscript address hyperspectral reflectance data treatment. It proposes a protocol to analyse data by combining spectral angle mapper and hyperspectral data manipulation.

As stated in the manuscript, hyperspectral imaging datasets are difficult to handle because of size, artefacts of the spectra, expertise needed. The proposal of a robust algorithm available for all users is of great interest. Moreover, this subject seems to me particularly relevant for a video.

The manuscript question is clear, and the author aware of the limits of their protocol. As far as I can judge, the manuscript is written in correct English.

We thank you for your opinion about our work.

I have to mention that I am a little bit confused by the format of the manuscript I have to review. The subject seems of interest, but in the absence of video, it is not easy to judge the relevance of the examples and illustrations. In consequence I have no specific comments.

In our intent, the figures should illustrate the main issues dealing with the acquisition and analysis of a hyperspectral reflectance data set: the care needed for setting the experimental set-up (figure 2 and 3); the attention to the quality of the obtained data (figure 4); the illustration of the range of possibility offered by the protocol to investigate the data cubes (figure 5, 6 and 7) and finally the presentation of an innovative way to shed light over the information contained in the data-set (figure 7 and 8).

We hope that the video would confirm you in your opinion about our work.

Reviewer #4:

Manuscript Summary:

The manuscript describes the pipeline of an operative protocol to perform the analysis of hyperspectral reflectance imaging (RIS) dataset using the Spectral Angle Mapper (SAM) algorithm combined with user-based reference selection. There is nowadays an increasing need for the implementation of such dedicated protocol for the cultural heritage community, as there is an exponential growth of the amount of data being acquired at the surface of artworks using hyperspectral cameras, and a crucial need for improved automated data analyses. This work represents a first step toward the development of software dedicated to the analyses of RIS data from cultural heritage, and as such should be encouraged.

Major Concerns:

Other freely available software/GUI presenting similar possibilities in term of data management/analyses exist and are not mentioned in this paper, for example: SpectrononPro (https://resonon.com/software), and Hyper Tools GUI (https://www.hypertools.org/). It is of prime importance to reference those and compare the protocol developed in this paper to what is already available to users.

Thank you for your suggestion. The rationale behind the development of the protocol is not to provide the users with the best performant data handling or analysis method but rather to move the attention on the possibility to consider data manipulation as a concrete chance to better understand the sample under investigation. Since the characteristics of the protocol were initially developed on the specific case of the data from Quarto Stato and then extended to be applied to hyper-spectral data from any type of painting, we preferred to create a home-made dedicated software rather than search the most suitable in the literature and adjust it to our needs. In shed of these considerations, we think that a comparison with existing tools is not mandatory. Anyway, we understood your point and we add a justification for our choice at the end of the introduction section.

Moreover, we had a look at the two solutions you suggested (We also cited them in text). We find these solutions very interesting and this is our feeling about them.

SpectrononPro: it offers the possibility to browse through the image of the monitored area in order to identify the main spectra within the cube (what we can call end members). It allows to evaluate similarity maps to determine the spatial distribution of the spectra in the image; it seems very effective for feature classification and object recognition especially if you have the same shape multiple times. However, it is a little tricky for our case: we cannot find an easy way to evaluate the similarity map between a spectrum from a data cube A with the spectra of a data cube B; a key feature when you analyse a painting where each area can be peculiar. Maybe this comparison would be possible by developing a home-made plug-in, but, for reasons of time, we did not explore the SpectrononPro coding possibilities. Moreover, the manipulation of the data is not allowed; you can visualize the portion of the spectrum you need but you cannot, at least in our experience, exclude one or more wavelength ranges from the analysis. Finally, we cannot find a way to import a single spectrum from our databases (i.e. from data set acquired in the past in our lab or employing an instrument different from the IQ camera). In summary, for our case study SpectrononPro seems to be useful for a preliminary inspection, as well as the software released with the instrument, but not for the analysis we propose in the manuscript.

Hypertools: we judge Hypertools as a more complete handle and analysis tool with respect to SpectrononPro. Hypertools allows to import more cubes simultaneously, it offers the possibility to pre-process the data and in the GUI the user can also find something that it is similar to what we called data manipulation; in facts both the spectral and the spatial range to be analysed can be set by the user; however while the spatial selection appear very well developed the spectral selection appears quite basic; the user can select a range of wavelengths but he cannot merge different wavelength ranges nor arrange the wavelengths following a defined criterion (option that covers a key role in our protocol). The user can select the analysis method within a wide range of possibility and save the obtained results; this fact, in particular, makes Hypertools an effective tool for hyper-data exploration and analysis but in our, even short, experience the learning curve for use it is quite steep and it is generally necessary to refer often to the tutorials and videos on youtube. Besides the impossibility to rearrange the spectra following a criterion, Hypertools present another main limitation in our opinion: it works with the variables of the matlab workspace; this means that it does not directly manage the I/O (i.e. it does not import the raw data from instruments or automatically save the results) nor provides for the possibility of mixing data with different spectral resolution (we did not find the way to do something similar in the GUI). In summary Hypertools is a very interesting solution for hyperspectral data manage and analysis, and it could be adjusted to satisfy most of our needs; however the need to pre-process the data to make it appropriate for Hypertools together with the fact that an extensive training of the user is needed, confirm us in the idea that developing our own protocol through a precise set of dedicated functions still makes sense.

It is also important to add a section in the introduction about SAM, and a justification of why this method solely was considered for classification purposes.

We introduced the fundamentals of SAM in the text.

The SAM has been chosen because it is widely used (also the Specim IQ camera is provided with a SAM calculation protocol), it is easy to be understood (indeed it is enough to be familiar with the concept of the inner angle between two vectors) and, finally, it allows to evaluate the spatial distribution of the considered spectra; i.e. to understand how the pigments have been organized by Pellizza Da Volpedo during the realization of Quarto Stato.

Minor Concerns:

Language could be overall improved.

We tried to improve the language of the text