G Model
CULHER-3198; No. of Pages 11
ARTICLE IN PRESS
Journal of Cultural Heritage xxx (2017) xxx–xxx

Available online at
ScienceDirect
www.sciencedirect.com
Elsevier Masson France
EM|consulte
www.em-consulte.com

Original article

Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication

Adam Polak [a,b,*], Timothy Kelman [a], Paul Murray [a], Stephen Marshall [a], David J.M. Stothard [b], Nicholas Eastaugh [c], Francis Eastaugh [c]

[a] Centre for Signal & Image Processing, University of Strathclyde, 204 George Street, Glasgow G1 1XW, UK
[b] Fraunhofer Centre for Applied Photonics, Fraunhofer UK Research Ltd, 99 George Street, Glasgow G1 1RD, UK
[c] Art Analysis and Research Ltd, 162-164 Abbey Street, London SE1 2AN, UK

INFO ARTICLE

Historique de l'article :
Reçu le 18 mai 2016
Accepté le 30 janvier 2017
Disponible sur Internet le xxx

Keywords :
Hyperspectral imaging (HSI)
Infrared
Artwork authentication
Support vector machine (SVM)

ABSTRACT

In recent years various scientific practices have been adapted to the artwork analysis process. Although a set of techniques is available for art historians and scientists, there is a constant need for rapid and non-destructive methods to empower the art authentication process. In this paper hyperspectral imaging combined with signal processing and classification techniques are proposed as a tool to enhance the process for identification of art forgeries. Using bespoke paintings designed for this work, a spectral library of selected pigments was established and the viability of training and the application of classification techniques based on this data was demonstrated. Using these techniques for the analysis of actual forged paintings resulted in the identification of anachronistic paint, confirming the falsity of the artwork. This paper demonstrates the applicability of infrared (IR) hyperspectral imaging for artwork authentication.

1. Introduction

According to a recent studies, in 2014 the global art market reached its highest ever-recorded level of just over €51 billion worldwide [1]. This represents a 7% year-on-year increase from €47.4 billion recorded in total sales of art and antiques in 2013, consisting of more than 36 million transactions [2]. The vast majority of these high value dealings were made without scientific or forensic testing to assure the authenticity of the traded objects. Non-scientific art expertise – known in the art world as connoisseurship – is a common practice to assess the authenticity. Nevertheless, an experienced specialist can only evaluate a limited amount of the artwork and when not supported by additional scientific tests, that evaluation is subjective and as such it is not infallible [3,4]. Services using scientific approaches to determine the authenticity of artworks are available; however, these can have perceived issues, including the time involved and the need to remove sample material for a number of the techniques [3]. There is consequently a need for efficient, portable and cost effective non-destructive methods of art analysis to serve a broader range of the market. In some cases, due to the high value and unique nature of the objects, the paint sampling required by certain types of examinations may also be restricted. Non-destructive tests provide the possibility to use complementary techniques and obtain more information from the same sample. Several such methods, for instance X-ray fluorescence and FTIR (Fourier Transform Infrared) or Raman spectroscopy, exist and are applicable for studying artwork [5,6]. Although these methods are commonly used for scientific art investigation as well as for some other applications, there is still a need for new, non-invasive techniques that could extend the amount of information obtained from the artwork analyses and limit the number of invasive testing required. In this research, Hyperspectral Imaging (HSI) combined with chemometrics algorithms is proposed as a novel, non-invasive analysis method for classification and mapping of paints and pigments. The aim is that these tools will serve as an aid for artwork evaluation and specifically, the identification of counterfeits.

In recent years Hyperspectral Imaging has undergone significant development. There is an increasing amount of camera technologies that, with different configurations, provide many ways to obtain hyperspectral data over several spectral ranges. This emerging technology is rapidly finding applications in different fields, including pharmaceuticals [7], agriculture [8] and food quality control [9–12], as well the art world, for material identification and mapping of the works of art [13–18]. To date, most

* Corresponding author at: Centre for Signal & Image Processing, University of Strathclyde, 204 George Street, Glasgow G1 1XW, UK.
Adresse e-mail : adam.polak@strath.ac.uk (A. Polak).

http://dx.doi.org/10.1016/j.culher.2017.01.013

Pour citer cet article : A. Polak, et al., Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication, Journal of Cultural Heritage (2017), http://dx.doi.org/10.1016/j.culher.2017.01.013

applications of hyperspectral (and multispectral) technology are for the restoration and conservation of paintings [19–22]. Pigment analysis provided by HSI systems coupled with dedicated classification algorithms allows the identification of "restored zones" in the painting and differentiates these from significant areas of the original painting which the system then suggests for pigment analysis [18,19,23,24]. The use of HSI in the infrared spectral range has also helped to reveal features of artists' techniques such as their preparatory drawing [22,25]. Due to the very broad range of wavelengths available for hyperspectral systems, the transmittance and reflectance response of different layers of paintings and drawings is frequently observed during data analysis. When material that is transparent at a specific wavelength range (but opaque at others) covers material that is reflective within the same spectral range, the underlying material can be detected by the HSI system and is hence revealed in the data acquired. Empowered by signal processing techniques this facilitates a detailed study of the artwork creation process and enables identification of the materials used [26]. Spectral selectivity of HSI data was also used on various occasions to analyse texts of historic value [27–29]. Identification of pigments and inks facilitated by HSI has also been used to aid in dating of manuscripts [27]. Furthermore, HSI technology empowers the recovery of erased and overwritten scripts as well as allowing the determination of appropriate bands for monitoring laser and non-laser cleaning processes [28].

Alongside the aforementioned benefits for art conservation, the potential of hyperspectral imaging in forgery detection has also been recognised. The application of HSI to address the challenge of forensic analysis of documents was successfully applied in the past [30–32]. Classification of different inks after obliteration of the text and the 'crossing lines problem' [30] were studied and analysed with chemometrics-based tools. These techniques applied to HSI data have had a significant impact on forgery recognition and provided objective results compared to traditional visual inspection based judgments [30]. Other forensic applications of HSI were also reported, such as fingerprint detection [33] and for blood stain dating at crime scenes [34].

The implementation of hyperspectral imaging in the aforementioned applications gave access to the rich dataset, however the conclusions drawn from this data were based on subsequent signal processing making it difficult to use by non-experts. In some cases the use of pure spectroscopic techniques achieved sufficient results [25], while in others, more advanced algorithms were applied in order to analyse the data [23,24,26]. It was also recognised that full diagnostic potential of HSI may be improved by implementation of robust data processing algorithms [28].

It is clear that HSI technology combined with advanced signal processing techniques have already found various application in the art world. However, to date these have focused on supporting various aspects of conservation and have allowed researchers to better understand paintings by allowing them to observe materials below the surface of the completed work. In this paper, we illustrate a novel combination of near- and mid-infrared hyperspectral imaging with state-of-the-art signal processing algorithms and background information from experts in the field of art analysis to provide HSI data based classification of paints and artwork for the purposes of authentication. As long as near infrared range was reached with widely known HSI technology, access to the mid-IR region was granted by the novel application of an active, laser-based mid-IR Imager. Although similar wavelength range was already explored with a passive system [35], to the authors' knowledge, our work presents the first ever application of this active device for the artwork analysis. This text demonstrates hyperspectral imaging empowered by automated paint classification techniques as a non-invasive method supporting the identification of counterfeit paintings. Our work was divided into two parts: (1) algorithms were developed using bespoke paintings which were created for this study and imaged in a well-controlled environment under laboratory conditions and (2) the techniques developed were applied to hyperspectral images of paintings held by the Berlin Landeskriminalamt which of comprised known and suspected forgeries, including, for the first time analysed with HSI, paintings from the infamous Beltracchi case [36–40]. This paper maintains this dual structure and focuses initially on describing system development and testing before providing details and the results achieved during this work. Some aspects of this study were also described in [41] however that text focuses on the difficulties of data acquisition and methods for overcoming these problems. Here, the focus is on the data analysis and processing as well as the application.

2. Materials and methods

2.1. Hyperspectral equipment

Applications of HSI systems operating in the visible–near-infrared (Vis–NIR) spectrum (400–1000 nm) have already been presented in the literature and tend to focus on performing and supporting various tasks including spectral characterisation of pigments [17,18,20,23,26–28]. InGaAs detector based hyperspectral imagers are also reaching further into near-infrared region (900–1700 nm, and in some cases extended up to 2500 nm) and these also have found application in the study of artworks [16–18,25]. However, relevant literature describing the use of sensors operating in longer wavelengths, approaching up to 4000 nm, which are known to contain rich spectral information and useful chemometric descriptors is not so readily available. In this paper, we therefore use two hyperspectral imaging systems operating in different (but overlapping) regions of the infrared portion of the electromagnetic spectrum. The choice of these two systems allowed us to study the impact of the image acquisition techniques and illumination methods [41] on the performance of our proposed signal and image processing techniques designed to automatically analyse the near- and mid-infrared range data. This work is driven by the motivation that in addition to the colour information contained in the visible spectrum, often sufficient to identify various pigments, a range of paint types (including pigments, binders and solvents) also have spectral features in the longer wavelengths. Hence this study is focussed on exploring these for the accurate discrimination of paints. It should be noted however, that while the intention of this study is to explore the usefulness of these longer wavelengths, many pigments can be discriminated using the Vis–NIR region and this could be beneficial for the final application of this technology by art scientists. As such, we discuss this topic further in Section 4.1.

The hyperspectral imaging systems which were employed during this study were: an active, laser based, mid-infrared hyperspectral imager (Firefly IR Imager, M Squared Lasers – see Fig. 1a) and a passive hyperspectral camera operating in the near-infrared wavelength range (Red Eye 1.7, inno-spec GmbH – see Fig. 1b).

While the near-IR range covered by the passive system (from 900 nm up to 1700 nm) is becoming more common and can be captured by various systems, devices operating in the infrared bandwidth beyond 3000 nm are still quite rare. The Firefly IR Imager is based on Optical Parametric Oscillator (OPO) technology that, with its inherent narrow spectral linewidth and wavelength tunability makes a foundation for a new class of hyperspectral imaging technology that is able to sense radiation beyond 3000 nm. It achieves this by converting radiation from a fixed frequency near-infrared laser source (1064 nm) into broadly tunable radiation in the mid-IR portion of the spectrum (2500–3750 nm) where compounds contained in paints exhibit distinct optical absorption.

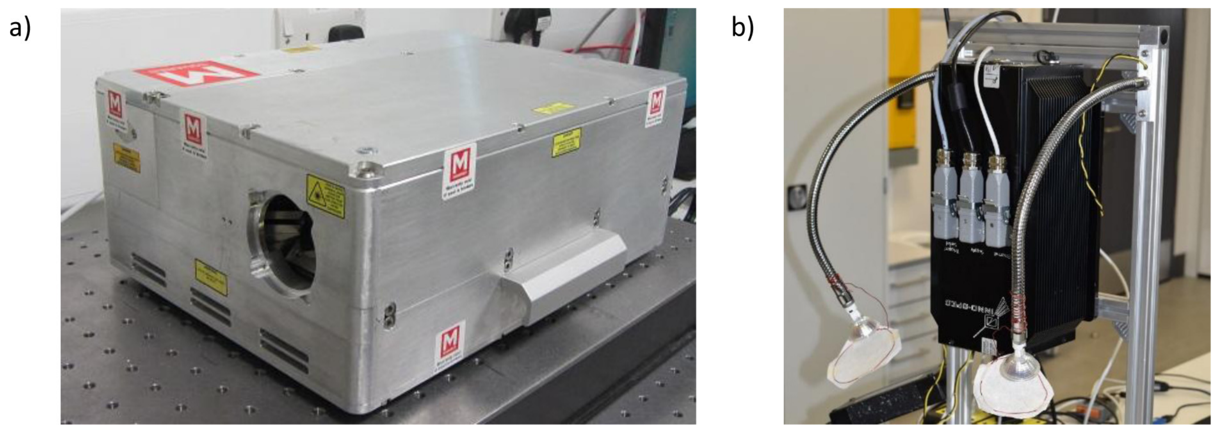
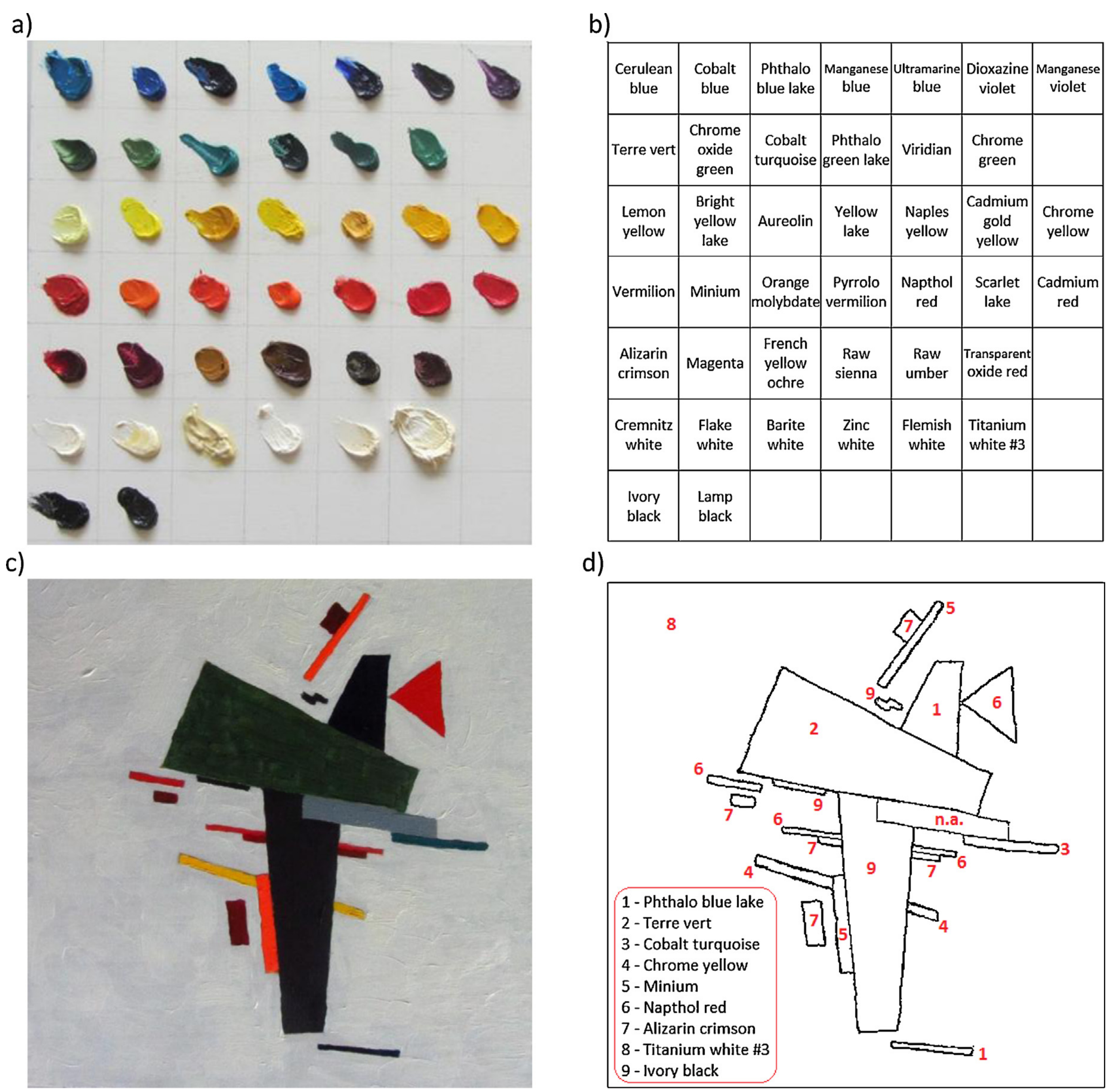


Fig. 1. Hyperspectral systems used during the project; (a) Firefly IR Imager; (b) Red Eye 1.7.



Cerulean blue	Cobalt blue	Phthalo blue lake	Manganese blue	Ultramarine blue	Dioxazine violet	Manganese violet
Terre vert	Chrome oxide green	Cobalt turquoise	Phthalo green lake	Viridian	Chrome green	
Lemon yellow	Bright yellow lake	Aureolin	Yellow lake	Naples yellow	Cadmium gold yellow	Chrome yellow
Vermilion	Minium	Orange molybdate	Pyrrolo vermilion	Napthol red	Scarlet lake	Cadmium red
Alizarin crimson	Magenta	French yellow ochre	Raw sienna	Raw umber	Transparent oxide red	
Cremnitz white	Flake white	Barite white	Zinc white	Flemish white	Titanium white #3	
Ivory black	Lamp black					

Fig. 2. Illustration of paintings used during the lab stage of the project and their ground truth description; (a) the pigment grid canvas serving as a training data; (b) description of the pigments used in the grid canvas; (c) Pastiche of "Untitled (Suprematist Composition)," by Kazimir Malevich as testing painting; (d) ground truth data of pigments used for creation of the Kazimir Malevich pastiche [41].

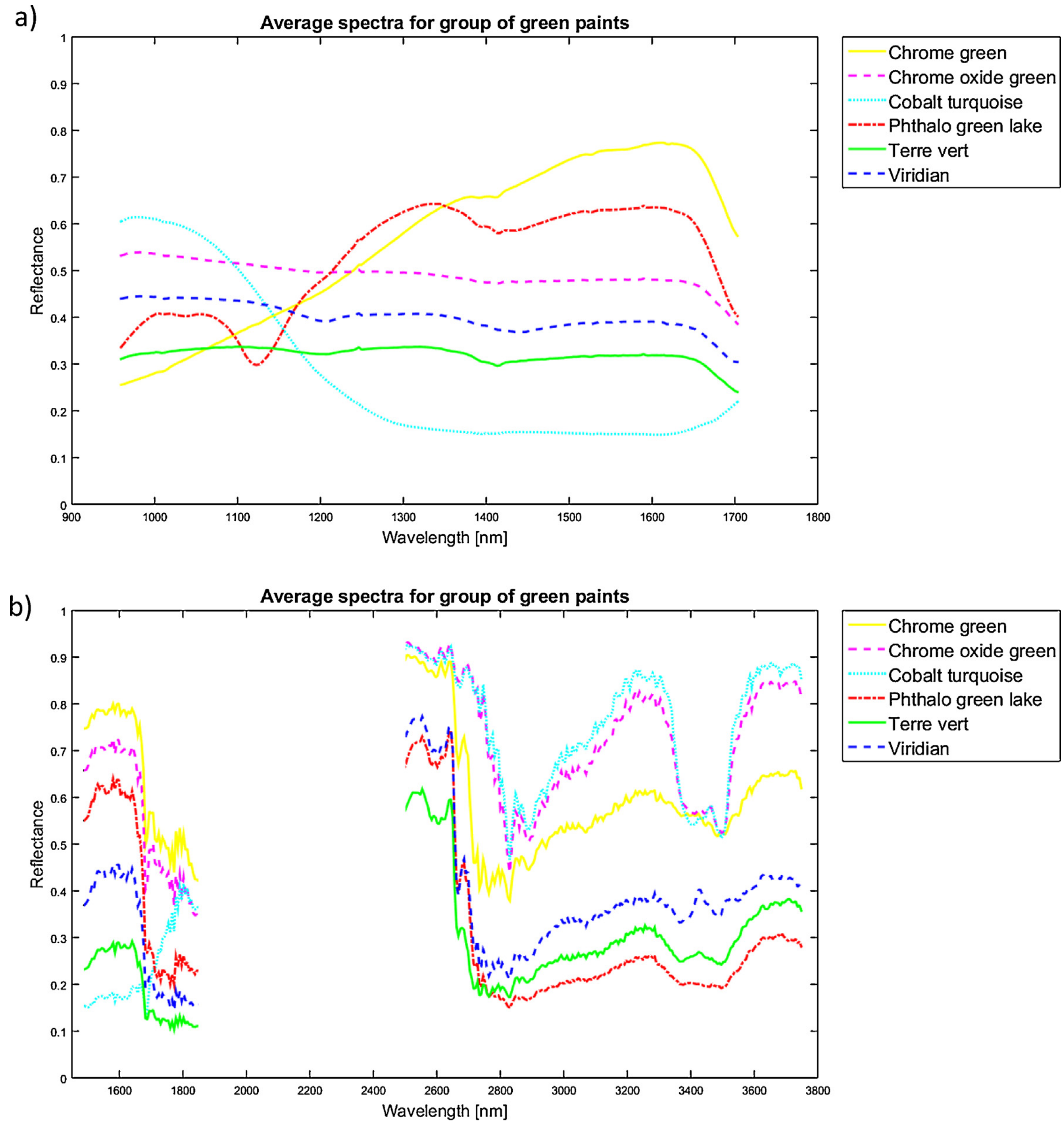


Fig. 3. Average spectral response of all the green coloured pigments acquired by passive (a) and active (b) system.

Thanks to the OPO technology, the Firefly IR Imager also features a narrow spectral band in a near-IR range (1490–1850 nm) thus providing a good link between both imaging systems used in this work.

2.2. Building and validating a spectral library using bespoke paintings

Two different paintings were prepared specifically to allow the construction of a "spectral library" from HSI datasets containing only known paints at specific locations. Both paintings were used for algorithm development and validation after the spectral library was created. One type of painting used mainly for development was constructed in the form of a grid of 41 oil-based paints based on a selection representative of the materials commonly used by artists during the 20th century (see Fig. 2a) [42]. Exemplar paints were purchased from a number of specialist artists' colourmen with particular reference to those offering ranges that include ostensibly "historically appropriate" materials (Michael Harding; Rublev; Blockx). The full description of all paints contained by this matrix can be found in Fig. 2b and Appendix 1. All paints used were also analysed using other instrumental techniques such as scanning electron microscopy-energy dispersive X-ray spectrometry ("SEM-EDX"), FTIR spectroscopy and Raman microscopy; identifications were made with reference to spectral libraries derived from the Pigmentum Project collection of historical pigments [42].

Two of these "grid canvases" as described above were used for signal processing algorithm development and also served as training data for classifications of other unseen paintings believed to contain at least one or more of these pigments. A second style of canvas, used for algorithm validation, was a pastiche of a Suprematist work by Kazimir Malevich (see Fig. 2c). This bespoke painting was created using a selective subset of the paints contained in the grid canvas (Fig. 2a), and was accompanied by labelled "ground-truth" data describing each area painted with different paint (see Fig. 2d).

Examination of the second painting provided controlled conditions for validation of the algorithms developed for automatic paint recognition. These two analysed paintings (grid canvas and

Pour citer cet article : A. Polak, et al., Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication, Journal of Cultural Heritage (2017), http://dx.doi.org/10.1016/j.culher.2017.01.013

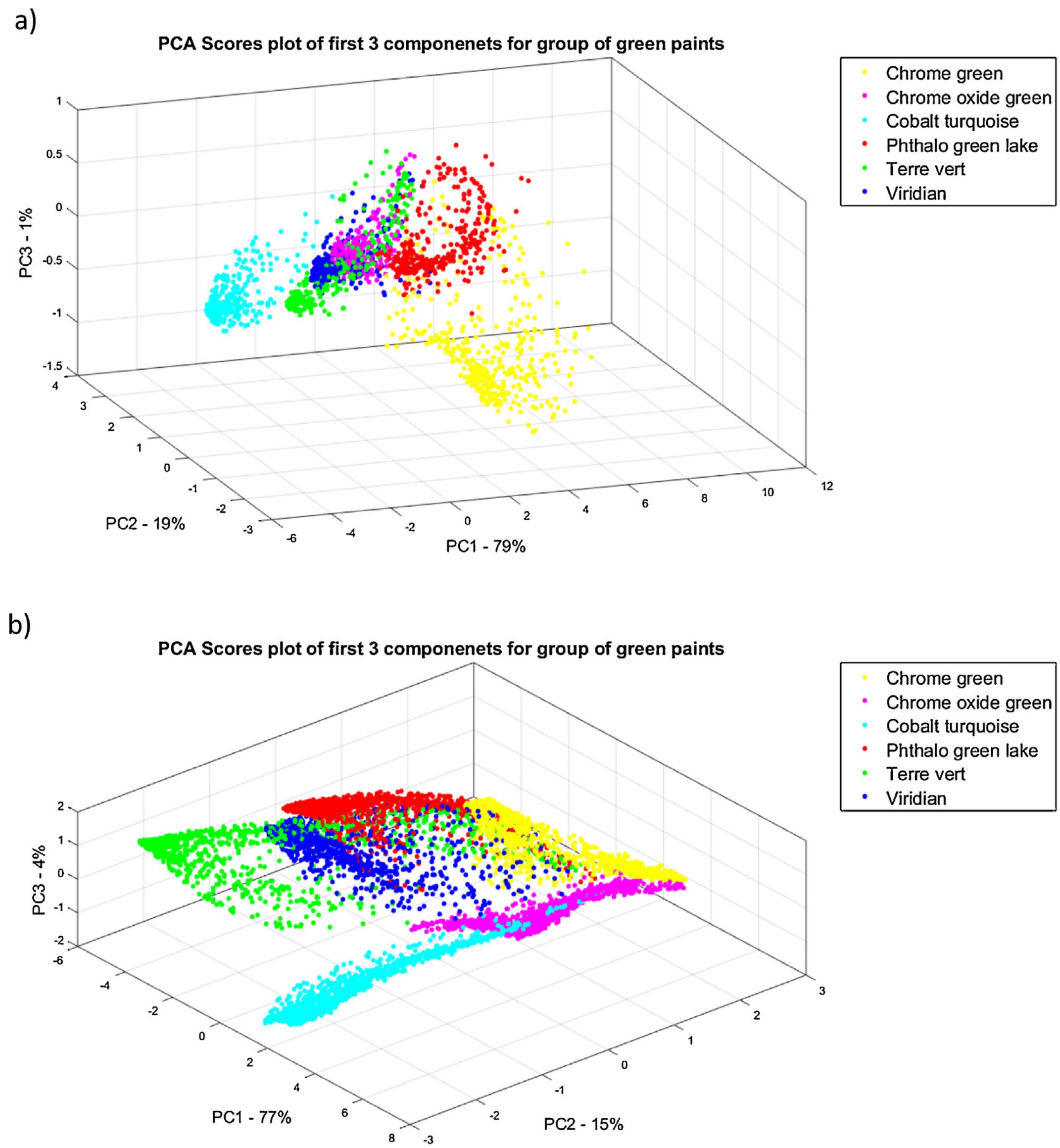


Fig. 4. PCA scores plot of first three principal components for the group of green paints imaged with passive (a) and active (b) system.

pastiche) constituted not only a very well controlled but also a realistic data set of paints which are commonly used in 20th century artwork upon which the methods developed in this work could be assessed.

2.3. Data acquisition and pre-processing

Due to significantly different mechanisms driving the two hyperspectral imagers used during this project, data acquisition and its preparation for analysis varied considerably between each case [41].

2.3.1. Passive system

The passive Red Eye 1.7 system acquires HSI data cubes using a pushbroom technique [43] and requires relative movement of the object in front of the detector. A Zolix KSA 11-200S4N linear translation stage was used to scan the paintings. Illumination was provided by off-the-shelf 12 V DC halogen reflectors. The halogen lamps, as an incandescent light source, emit not only a portion of energy in the visible band of the electromagnetic spectrum, but also provide excellent illumination in the near-IR range as required by this system [44]. During each data acquisition run, reflectance calibration was also performed to ensure that background spectral responses of the instrument and illumination, as well as the 'dark' current of the camera were accounted for in the data set and therefore do not affect the results of any subsequent analysis. The relative reflectance for the raw images can be calculated using:

$$R = \frac{I_0 - D}{W - D} \quad (1)$$

where R is the relative reflectance image, I_0 is the raw reflectance image, D is the dark reference image, and W is the white reference image [43]. The spectral background W was obtained by scanning a white reference tile made from Spectralon – a material of high Lambertian reflection over its reflective spectral range of 250–2500 nm [45]. The dark reference D was captured by fully obscuring the camera objective using an opaque black cap. All hyperspectral images of the paintings were prepared in the same way using Eq. (1) prior to the analysis and no additional pre-processing was applied.

2.3.2. Active system

The active Firefly IR Imager acquires HSI data using active laser illumination which scans over its working spectral range and no external illumination is required. This imager is equipped with a built-in scanner based on two oscillating mirrors providing spatial scanning of the laser beam across the target object and therefore employing a whiskbroom scanning technique [43]. To minimise the effect of whiskbroom scanning artefacts (spatial distortion and intensity variation), and to increase spatial resolution of the image, all analysed paintings were scanned in sections which were then stitched together to recreate whole spatial area of the painting in one hyperspectral data cube [41]. The size of the painting sections were chosen as 50 × 75 mm (this size was derived empirically) and a full paint grid canvas was reproduced from 9 such sections, while the slightly smaller Malevich pastiche consisted of 6 sections. Due to the nature of the hardware (complexity of detector settings), the reflectance calibration was not possible for this equipment. To rescale the spectral data to the reflectance range [0–1], the 8-bit intensity data acquired on each wavelength was divided by the maximum value (255). Thanks to the continuous spectral tunability of the laser source, the acquisition step may be arbitrarily set by the operator, with the hardware limitation of 0.1 nm. However, due to the finite linewidth of the laser source (approx. 5 cm^{-1}) and shape of spectral features of imaged paints, 6 nm spectral resolution was chosen, providing 61 spectral image bands in the near-IR (1490–1850 nm) and 209 image bands in the mid-IR region (2500–3750 nm). After completing the data stitching and rescaling, no further pre-processing was applied.

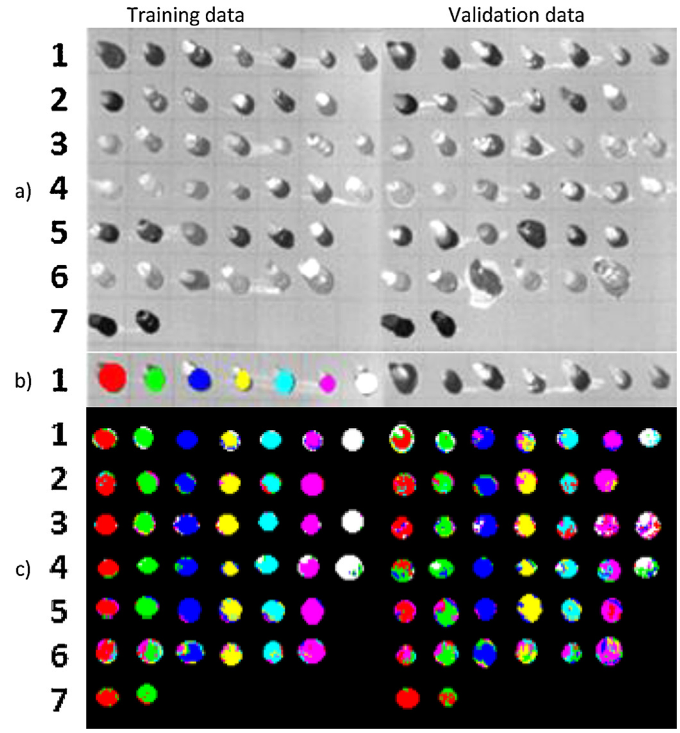


Fig. 5. Illustration of the classification algorithm validation based on Red Eye 1.7 data; (a) intensity image on one wavelength for two grid canvas; (b) single row with assigned colour labels to each paint; (c) classification result of the training data set (on the left side) and validation data set (on the right).

3. Algorithm development and feature extraction

After the pre-treatment to normalise and calibrate the data (see Section 2.3), subsequent processing was the same for both data sets and an algorithm was designed to facilitate hyperspectral analysis of the artwork. Since this work aimed at the identification of different paints, the main aspect of algorithm development was focused on the application of robust statistical classification techniques. Both supervised (guided by human provided training data) and unsupervised (fully based on software analysis of the image) techniques were considered [46]. Since the objective of work is the detection of counterfeit paintings by the classification of known paints, supervised classification was chosen as the most suitable for this application. The grid canvas presented in Fig. 2a) was used to build a spectral library of selected pigments and this served as training data for the algorithm. Fig. 3 illustrates the average spectra (acquired by both systems) of the group of green paints available on the grid canvas that is shown in Fig. 2. All the average spectra from the 41 paints in the library are presented in Appendices 2 and 3.

Although this figure illustrates the averaged spectral response from the entire paint regions, the algorithm is trained to recognise all variations of this response resulting from the uneven paint surfaces when imaged. As it is expected that during the analysis of artwork the captured data will also contain various artefacts of imaging process, all these variations were also allowed in the training set classes. The training data therefore included examples of specular reflections and intensity variations coming from the three-dimensional structure of the paint blobs and their thickness. A Principal Component Analysis (PCA) was performed to assess the quality of this training set and a PCA scores plot showing the first three principal components, explaining over 95% of variance in the data, was plotted. The examples of these plots for the group of green paints acquired by both imaging systems are shown in Fig. 4. For completeness all PCA scores plots for the full acquired library are shown in Appendices 4 and 5.

Although there is a wide range of algorithms which could be chosen to perform multivariate analysis using supervised classification, a Support Vector Machine (SVM) was identified as the most suitable method for this project due to its robustness and consistent classification accuracy [47–49]. SVMs consist of a family of learning algorithms that can be used for data classification and demonstrates very good performance in many applications [47,48,50]. For this reason it is one of the most popular machine learning algorithms that is often used to solve classification problems. The specific algorithm used for the paint classification in this study is contained in the LIBSVM package [51], which uses a C-SVM type of classification. As the SVM is a binary classifier, the multiclass problem was approached using a 'one-against-one' technique. The quality of the model was also evaluated using Cross Validation Accuracy scores acquired during model training and final Classification Accuracy [51].

In order to facilitate the construction of a spectral library and algorithm development, two of the "grid canvases" (see Section 2.2 and Fig. 2) were imaged side by side, and one was used as training data set while the other as validation set (see Fig. 5a). This approach alleviates the need for k-fold cross validation, as this introduces new, 'unseen' data for the classifier. Regions of interest (ROI) – subsets of the image manually selected to contain a group of pixels corresponding to a single paint – were defined for all 41 paint samples of the grid canvas. Since each row in the canvas represents one group of colours (see Fig. 2a), each row of the grid, labelled 1–7, was considered separately in the HSI data for algorithm development. Each paint is allocated a label denoted by colours in Fig. 5.

The leftmost painting (Fig. 5a) was used as training data and with use of ROIs, each paint of the selected row was assigned a colour label for classification (see Fig. 5b). During this project it was chosen not to compress the spectral dimension and instead the full spectral profile was used as the data features for algorithm training. After training, the classification was performed on both paintings and the whole process was repeated for each row 1–7 in turn. Satisfactory classification results were obtained for both – the

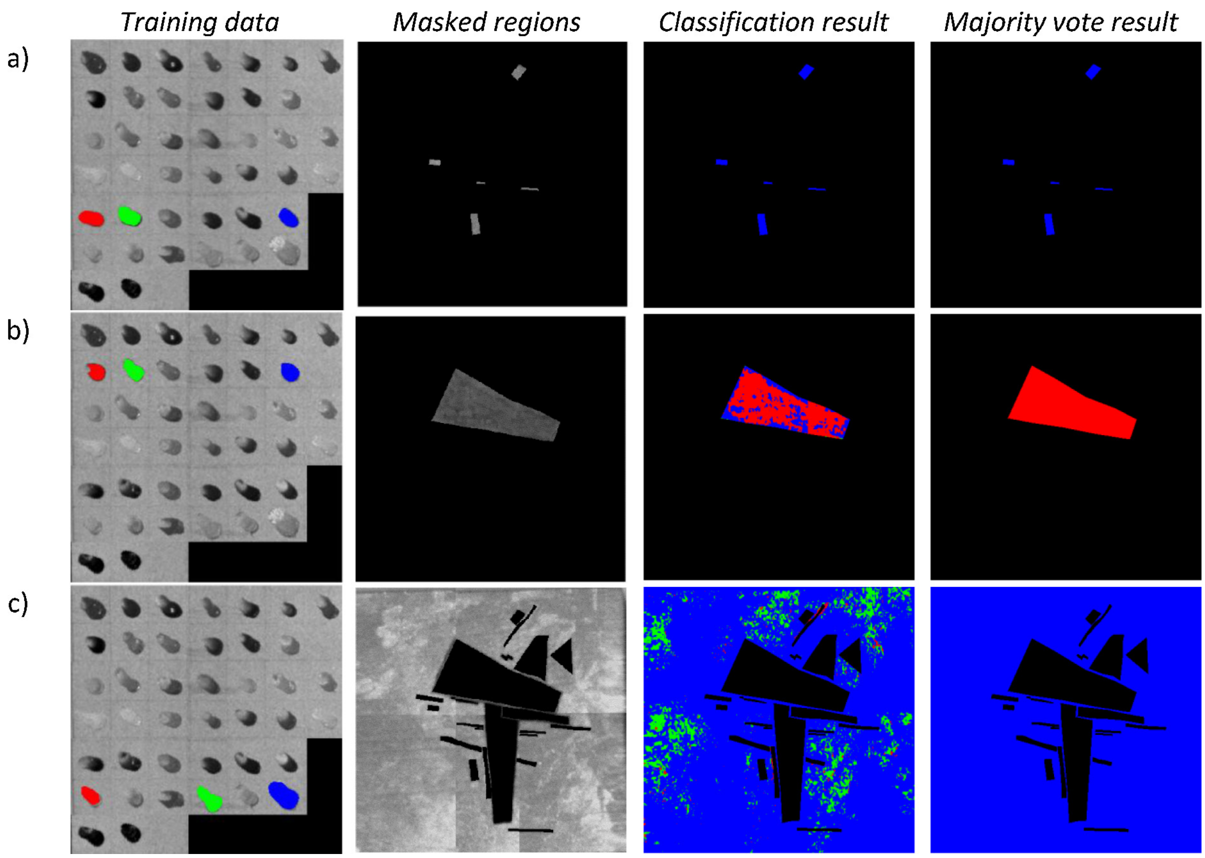


Fig. 6. Illustration of the analysis approach for pigment classification on tested painting. For each colour on the painting subset of training paints was chosen (first column – Training data) and classification was performed on the masked area of the painting corresponding to this colour (second column – Masked regions). Classification resulted in per-pixel classification of the selected area (third column – Classification result) and majority vote was drawn for these regions resulting in selection of one pigment corresponding to one colour (fourth column – majority vote result). This figure shows three examples of this approach for the regions of brown (a), green (b) and white (c) paint (demonstration based on Firefly IR Imager data).

dataset that classifier was trained on, as well as the new, previously unseen data set to which the classifier was applied. A graphical illustration of the combined result for each row when applying the proposed technique to each row separately is presented in Fig. 5c.

4. Classification

4.1. Candidate paint selection

In this paper, a dual stage classification process is proposed. To limit the amount of training data built into the classification model to recognise an individual paint, the first stage of classification aimed at choosing a subset of 3 potential paints from the hyperspectral image of the "grid canvas". This method not only accelerated the classification process and reduced computing power required for the problem by comparing only likely candidates, but also reduced the chance of misclassification. In practice, it does not make sense to attempt to recognise one single paint by referencing and comparing it with the entire spectral library of all paints available – especially those which are clearly a different colour. It should be noted that many pigments can be accurately distinguished by analysing the visible region of the spectrum and, for these, this primary classification may be sufficient. For all others, shortlisting likely candidates based on the colour and appearance is a practical solution to reduce computational overhead and the likelihood of error. In fact, while the presented concept has already demonstrated its potential with a training set of only 41 paints, with the expected extension of the library to contain hundreds of different pigments, a pre-selection such as that would be necessary for efficient performance of spectral data based classification. To date, the selection of candidate paints has been carried out manually based on paint colour, however an RGB or a visible range HSI system could also be used for this task and algorithms will be developed to perform this initial candidate selection step before the final spectral classification is made. Furthermore, based on the ground truth data, the described visual inspection always shortlisted the correct paint and therefore it was clearly a viable solution and one which could be easily performed – even by a non-expert user of the technology.

4.2. Accurate spectral classification

Fig. 6 provides examples of regions selected for classification and paint candidates chosen as training data for their classification. Due to the highly geometrical structure of the painting analysed in Fig. 6 (also shown in full and in colour in Fig. 2c), masks (binary images used to specify regions of the image for processing) were created for each painting area. Each structure on the painting was prepared with a single, unmixed paint. Therefore, to make demonstration of the classification results simpler, regions of the painting corresponding to a single colour were classified separately, with the remainder of the image being disabled from classification by application of the masks.

After the candidate pre-selection, a spectral classification stage using a SVM classifier and majority voting scheme was applied to assign the chosen paint region to one of the selected classes based on the spectral signature in each pixel of the masked region in the test painting. The classification was performed on a pixel-by-pixel basis, but with the prior knowledge that each section of the

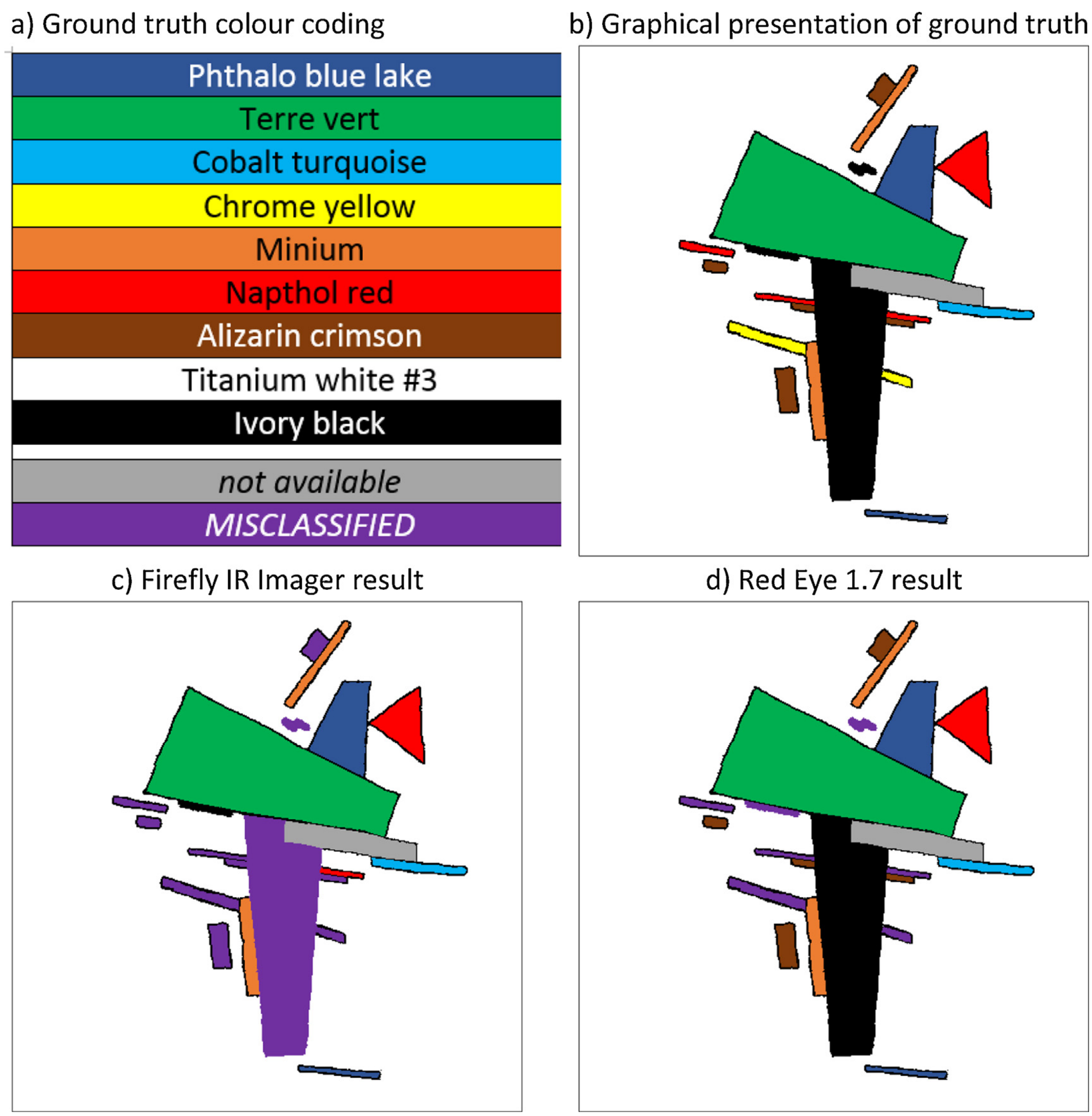


Fig. 7. Graphical illustration of class labelling for all the regions of the test paining; (a) the colour coding of the result demonstration; (b) Ground truth data illustration; (c) the result based on data from Firefly IR Imager; (d) the result based on data from Red Eye 1.7 system [41].

test painting was created with one type of the pigment – i.e. no mixtures or other impurities. Fig. 6 illustrates this approach and provides the classification and majority vote results for selected example regions. The colour of the label assigned to each pixel and subsequently each region of paint corresponds to the "class colour" assigned to the ROI selected from the training data set.

The approach shown in Fig. 6 was applied to all regions of the painting, assigning one class from the training set to each of them in turn. A total of 10 different paints were used to prepare the painting under study, while 9 of them were available in the "grid canvas" and were therefore contained within our spectral library. From this perspective, using the Firefly IR Image data it was possible to automatically classify 67% of pigments (6 out of 9) correctly, while the correctness ratio for data from the Red Eye 1.7 system reached 78% (7 out of 9) [41].

Fig. 7 shows the classification result for both imaging systems when compared with the ground truth data. It is clear that in some cases, if the classification of data from one device is wrong, it may be correct if the other is used, and vice versa. For example, one of the regions painted with Ivory black paint is always classified correctly by both systems, but it exists in a different region for both of them (see Fig. 7c and d). This demonstrates the complementary nature of the two devices and also shows that the full range of wavelengths considered is useful for discriminating different paints. Whilst the classification accuracies did not reach 100%, the potential of automated classification techniques based on the hyperspectral data has been clearly demonstrated.

5. Experimental results

Having demonstrated the proposed system's ability to recognise the paints based on spectral signatures in a well-controlled lab-based environment, this technology was applied to assess other paintings, suspected as or, by other means, already identified as forgeries. Thanks to the courtesy of the Berlin Police it was possible to image several paintings from their collection of forged paintings created by the infamous Wolfgang Beltracchi [38].

Due to the practicalities in transporting the equipment to Berlin, only the Red Eye 1.7 camera was used during this experiment. Data acquisition, pre-processing and analysis steps followed the approach described in Sections 2 and 3. The only difference was at the stage of classification since the analysed paintings did not feature the simple geometric properties of the paintings designed

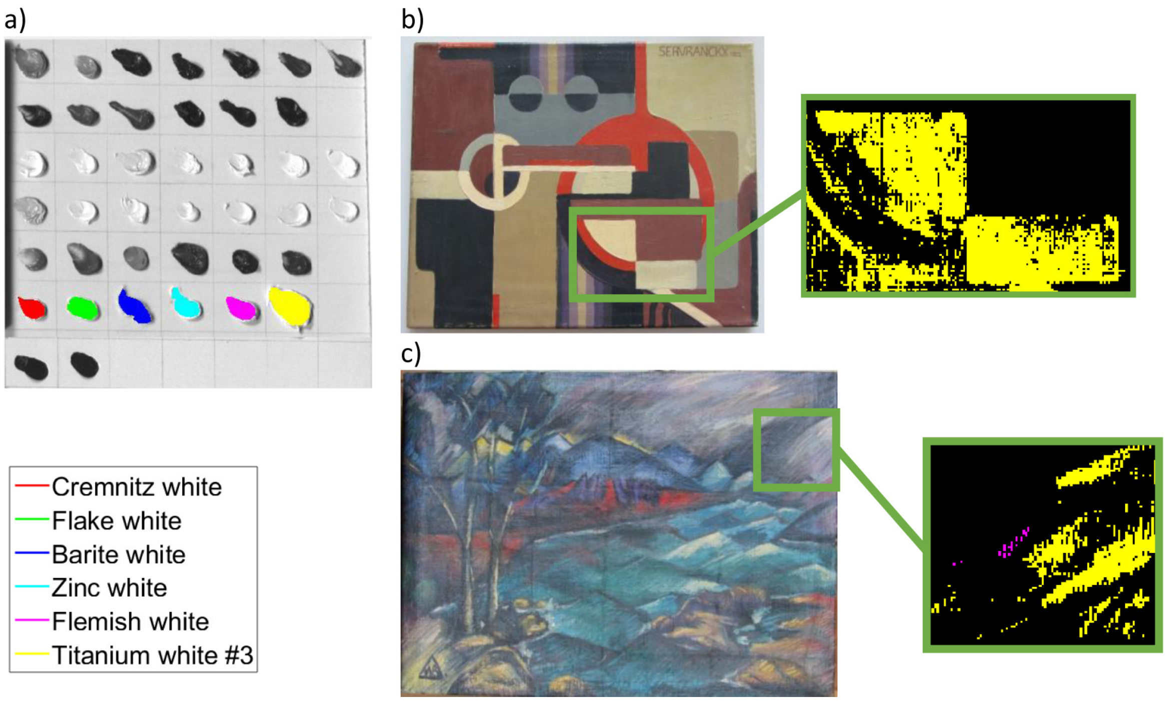


Fig. 8. Illustration of two forged paintings with indication of analysis region and classification result of selected colours; (a) intensity image of grid canvas with selection of paints used for training and the legend identifying the colour coded pigments; (b) painting described as Sevranckx with white/cream colour classification result; (c) painting referred as HM501 and white colour classification.

for this work as shown in previous sections. For this reason, the application of binary masks was not a feasible approach to allow individual regions of the paintings to be processed separately. Additionally, these paintings contained a variety of pigments, both in pure and mixed form. To overcome these challenges, the proposed techniques were applied to selected regions of two paintings, chosen based on other instrumental testing techniques (see Section 2.2), which contained "*Titanium white*" paint known in this case to be anachronistic. Fig. 8(b and c) shows the two selected paintings where the green box outlines the region chosen for the analysis alongside a magnification of the colour representation of the classification result.

Fig. 8(b) shows a painting referred to as 'Sevranckx' in which the white/cream regions have been identified as *Titanium White* in a separate invasive tests. The developed software was trained on all of the white/cream paints in the spectral library constructed in this study using pigments in the "grid canvas" painting (see Fig. 8a) and included: *Cremnitz white*, *Flake white*, *Barite white*, *Zinc white*, *Flemish white* and *Titanium white #3*. The right hand part of the figure shows the classification result for a small region of the painting. Yellow in this image corresponds to *Titanium White* in the spectral library. The black regions correspond to pixels in the image where no classification has been made by the system as it does not recognise the spectra of these pixels as belonging to any paint in the spectral library provided. Furthermore, no other classification labels/colours are visible in the image. The reason for this is that the classifier did not identify any other part of this section of the painting as containing any of the other white paints. From this result, it is clear that two of the white/cream regions have been correctly identified and this was validated by comparing with previously captured ground-truth information. A similar situation can be observed in Fig. 8c) demonstrating a painting referred to as 'HM501', where similar classification steps successfully identified *Titanium White* in agreement with the result of the aforementioned separate analysis. It may be noticed that several pixels were misclassified as Flemish white (marked with magenta in Fig. 8c). However, the number of pixels in error is very small and this is to be expected of any automated classification scheme. In both these test cases correct identification of *Titanium White* was very important for the art scientists, as this was the anachronistic pigment used in Beltracchi forgeries that, combined with other evidence, such as the use of another anachronistic paint, Phthalocyanine Green, and a very limited pool of re-used canvases and stretchers, confirmed the falsity of these paintings [36].

6. Conclusions

Hyperspectral Imaging combined with advanced signal processing techniques is a valid and potent technique which can be used as a tool to support the process of artwork authentication by identification and classification of pigments. In this paper we have presented our work in developing software based signal processing algorithms to facilitate feature extraction and classification of paints from hyperspectral images of bespoke and fraudulent artwork captured in the infrared spectral region. A wide range in the infrared was accessed by use of a conventional near-IR HSI camera as well as a novel laser based imager, extending the readily accessible bandwidth of this portion of the near-IR spectrum with a subset of the mid-IR region. Collaboration of art historians and signal processing specialists made it possible to create a small spectral library of pigments, based on the tailored grid canvas, and to apply the classification techniques to distinguish different paints used on a specially prepared painting. Thanks to unique access to an extremely high-profile dataset, developed algorithms also allowed us to test – for the first time with this technique – a set of known Beltracchi forged paintings. The classification results show moderate-high correct classification rate already on the first approach to this problem. In fact, up to 78% of pigments used in the bespoke test painting were correctly classified using the Red Eye 1.7 HSI imaging sensor. Ultimately, and perhaps most importantly an anachronistic paint – *Titanium White* – was identified from real forged paintings using our system in real world conditions

Pour citer cet article : A. Polak, et al., Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication, Journal of Cultural Heritage (2017), http://dx.doi.org/10.1016/j.culher.2017.01.013

using our separately acquired spectral library of pigments as reference.

This initial study demonstrates the effectiveness of hyperspectral imaging in combination with the image processing and classification techniques. However, it should be noted that some pigments were undistinguishable for the SVM classifier and further analysis would be required to determine which paints can be reliably identified with this method. Additionally, other spectral bands, feature extraction techniques and classification algorithms should be considered to further explore the effectiveness and improve the robustness of the final system. A robust spectral library of pigments is also crucial for successful classification. Various thicknesses of the paint, different base (e.g. wooden board or canvas), background painting surface and presence of varnish all may affect the spectral profile and should be considered for implementable system.

The application of scientific methods to reveal forgeries has recently gained significant public profile [4]. With increased demand for such analysis, the introduction of hyperspectral imaging can enhance the portfolio of available techniques and tools as a rapid and non-destructive method of artwork examination. Even if the results based solely on HSI data cannot provide full confirmation that a painting is a genuine or forged, current practices of art historians employ a combination of background knowledge and scientific data in the authentication process. By adding hyperspectral imaging to their already sophisticated toolkit, the analysis procedure may become faster, easier and more accessible to a larger portion of the market and the services should become more affordable as a result. Ultimately, the techniques proposed herein could limit the amount of destructive tests required for final validation and will increase the amount of analysed artwork thus making the whole procedure even more cost effective.

Acknowledgments

This study was carried out within the framework of an Intelligent Hyperspectral Imaging (INHERIt) Project financed by Innovate UK, Technology Strategy Board in collaboration between Department of Electronic & Electrical Engineering, University of Strathclyde (Glasgow, UK), M Squared Lasers (Glasgow, UK), Fraunhofer UK (Glasgow, UK) and Art Analysis & Research (London, UK). Authors wish to thank Berlin Landeskriminalamt for granting access to their paintings and supporting this study. This project was also supported by the EPSRC Centre for Doctoral Training in Applied Photonics, funded by the UK Engineering and Physical Sciences Research Council (grant EP/G037523/1) and by an Industrial Fellowship from the Royal Commission for the Exhibition of 1851.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.culher.2017.01.013.

Références

[1] C. McAndrew, TEFAF Art Market Report 2015, The European Fine Art Foundation (TEFAF), Helvoirt, 2015.
[2] C. McAndrew, TEFAF Art Market Report 2014: The Global Art Market, With a Focus on the US and China, The European Fine Art Foundation (TEFAF), Helvoirt, 2014.
[3] S. Hardach, BBC – Culture – The Surprising Secrets of Busting Art Forgeries, 2015, 19 October, Available at: http://www.bbc.com/culture/story/20151015-the-surprising-secrets-of-busting-art-forgeries (accessed 08.01.16).
[4] N. Eastaugh, J. Nadolny, Science for the Art Market: A Buyer's Guide. ArtBanc Market Intelligence, No. 7, 2007, April, pp. 33–37.
[5] A. Adriaens, COST Action G8: non-destructive analysis and testing of museum objects, in: Benefits of Non-Destructive Analytical Techniques for Conservation, Kalkara, 2004.
[6] G.D. Smith, J.F. Hamm, D.A. Kushel, C.E. Rogge, What's wrong with this picture? The technical analysis of a known forgery, in: Collaborative Endeavors in the Chemical Analysis of Art and Cultural Heritage Materials, American Chemical Society, 2012, pp. 1–21.
[7] Y. Roggo, A. Edmond, P. Chalus, M. Ulmschneider, Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms, Anal. Chim. Acta 535 (2005) 79–87.
[8] B.D. Jadhav, P.M. Patil, Hyperspectral remote sensing for agricultural management: a survey, Int. J. Comput. Appl. 106 (7) (2014) 38–43.
[9] G.K. Naganathan, L.M. Grimes, J. Subbiah, C.R. Calkins, A. Samal, G.E. Meyer, Visible/near-infrared hyperspectral imaging for beef tenderness prediction, Comput. Electron. Agric. 64 (2) (2008) 225–233.
[10] T. Qiao, J. Ren, C. Craigie, J. Zabalza, C. Maltin, S. Marshall, Comparison between near infrared spectroscopy and hyperspectral imaging in predicting beef eating quality, in: Hyperspectral Imaging and Applications Conference (HSI 2014), Coventry, 2014.
[11] T. Qiao, J. Ren, C. Craigie, J. Zabalza, C. Maltin, S. Marshall, Quantitative prediction of beef quality using visible and NIR spectroscopy with large data samples under industry conditions, J. Appl. Spectrosc. 82 (1) (2015).
[12] T. Qiao, J. Ren, J. Zabalza, S. Marshall, Prediction of lamb eating quality using hyperspectral imaging, in: OCM (Optical Characterization of Materials) 2015, Karlsruhe, 2015.
[13] S. Legrand, F. Vanmeert, G.V.d. Snickt, M. Alfeld, W.D. Nolf, J. Dik, K. Janssens, Examination of historical paintings by state-of-the-art hyperspectral imaging methods: from scanning infra-red spectroscopy to computed X-ray laminography, Herit. Sci. 2 (2014) 13–24.
[14] C. Fischer, I. Kakoulli, Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications, Stud. Conserv. 51 (Suppl. 1) (2006) 3–16.
[15] F. Rosi, C. Miliani, R. Braun, R. Harig, D. Sali, B.G. Brunetti, A. Sgamellotti, Noninvasive analysis of paintings by mid-infrared hyperspectral imaging, Angew. Chem. Int. Ed. 52 (20) (2013) 5258–5261.
[16] 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 (2013) 4838–4848.
[17] J.K. Delaney, J.G. Zeibel, M. Thoury, R. Littleton, M. Palmer, K.M. Morales, E.R.d.L. Rie, A. 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.
[18] J.K. Delaney, M. Thoury, J.G. Zeibel, P. Ricciardi, K.M. Morales, K.A. Dooley, Visible and infrared imaging spectroscopy of paintings and improved reflectography, Herit. Sci. 4 (6) (2016) 1–10.
[19] H. Liang, Advances in multispectral and hyperspectral imaging for archaeology and art conservation, Appl. Phys. A 106 (2) (2012) 309–323.
[20] A. Casini, C. Cucci, M. Picollo, L. Stefani, T. Vitorino, Creation of a hyper-spectral imaging reference database of red lake pigments, COSCH e-Bulletin (2015).
[21] A. Casini, F. Lotti, M. Picollo, L. Stefani, E. Buzzegoli, Image spectroscopy mapping technique for noninvasive analysis of paintings, Stud. Conserv. 44 (1) (1999) 39–48.
[22] M. Kubik, Hyperspectral imaging: a new technique for the non-invasive study of artworks, physical techniques in the study of art, Archaeol. Cult. Herit. 2 (2007) 199–259.
[23] F. Daniel, A. Mounier, J. Pérez-Arantegui, C. Pardos, N. Prieto-Taboada, S.F.-O.d. Vallejuelo, K. Castroc, Hyperspectral imaging applied to the analysis of Goya paintings in the Museum of Zaragoza (Spain), Microchem. J. 126 (2016) 113–120.
[24] S. Baronti, A. Casini, F. Lotti, S. Porcinai, Multispectral imaging system for the mapping of pigments in works of art by use of principal-component analysis, Appl. Opt. 37 (8) (1998) 1299–1309.
[25] C. Cucci, J.K. Delaney, M. Picollo, Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts, Acc. Chem. Res. 49 (10) (2016) 2070–2079.
[26] M. Attas, E. Cloutis, C. Collins, D. Goltz, C. Majzels, J.R. Mansfield, H.H. Mantsch, Near-infrared spectroscopic imaging in art conservation: investigation of drawing constituents, J. Cult. Herit. 4 (2003) 127–136.
[27] K. Melessanaki, V. Papadakis, C. Balas, D. Anglos, Laser induced breakdown spectroscopy and hyper-spectral imaging analysis of pigments on an illuminated manuscript, Spectrochim. Acta Part B 56 (2001) 2337–2346.
[28] C. Balas, V. Papadakis, N. Papadakis, A. Papadakis, E. Vazgiouraki, G. Themelis, A novel hyper-spectral imaging apparatus for the non-destructive analysis of objects of artistic and historic value, J. Cult. Herit. 4 (2003) 330–337.
[29] M.E. Klein, B.J. Aalderink, R. Padoan, G.d. Bruin, T.A. Steemers, Quantitative hyperspectral reflectance imaging, Sensors 8 (2008) 5576–5618.
[30] C.S. Silva, M.F. Pimentel, R.S. Honorato, C. Pasquini, J.M. Prats-Montalban, A. Ferrere, Near infrared hyperspectral imaging for forensic analysis of document forgery, Analyst 139 (2014) 5176–5184.
[31] Z. Luo, P. Shafait, A. Mian, Localized forgery detection in hyperspectral document images, in: 13th International Conference on Document Analysis and Recognition (ICDAR), Nancy, France, 2015.
[32] Z. Khan, F. Shafait, A. Mian, Towards automated hyperspectral document image analysis, in: Proceedings of the 2nd International Workshop on Automated

Forensic Handwriting Analysis: A Satellite Workshop of International Conference on Document Analysis and Recognition (ICDAR 2013), Washington, DC, USA, 2013.
[33] T. Chen, Z.D. Schultz, I.W. Levin, Infrared spectroscopic imaging of latent fingerprints and associated forensic evidence, Analyst 134 (2009) 1902–1904.
[34] G. Edelman, E. Gaston, T.v. Leeuwen, P. Cullen, M. Aalders, Hyperspectral imaging for non-contact analysis of forensic traces, Forensic Sci. Int. 223 (2012) 28–39.
[35] C. Daffara, D. Ambrosini, L. Pezzati, D. Paoletti, Thermal quasi-reflectography: a new imaging tool in art conservation, Opt. Express 20 (13) (2012) 14746–14753.
[36] C. Finn, The devil in the detail, Apollo: Int. Art Mag. CLXXIX (January (616)) (2014) 50–54.
[37] H. Keazor, T. Öcal (Eds.), Der Fall Beltracchi und die Folgen. Interdisziplinäre Fälschungenforschung heute, De Gruyter, Berlin, 2014.
[38] J. Hammer, The Greatest Fake-Art Scam in History? Vanity Fair, 2012, 10 October, Available at: http://www.vanityfair.com/culture/2012/10/wolfgang-beltracchi-helene-art-scam (accessed 22.03.16).
[39] S. Koldehoff, T. Timm, Falsche Bilder, Echtes Geld (Fake Paintings, Real Money), Verlag Galliani, Berlin, 2012.
[40] M. Sontheimer, A Cheerful Prisoner: Art Forger All Smiles After Guilty Plea Seals Deal, Der Spiegel, 2011, 27 October, http://www.spiegel.de/international/germany/0,1518,794454,00.html (accessed 18.03.16).
[41] A. Polak, T. Kelman, P. Murray, S. Marshall, D. Stothard, N. Eastaughc, F. Eastaugh, Use of infrared hyperspectral imaging as an aid for paint identification J. Spectral Imaging 5 (1) (2016) 1–10.
[42] N. Eastaugh, V. Walsh, T. Chaplin, R. Siddall, The Pigment Compendium, Elsevier, 2008.
[43] G. ElMasry, D.-W. Sun, Principles of hyperspectral imaging technology, in: Hyperspectral Imaging for Food Quality Analysis and Control, Elsevier Inc., London, 2010, pp. 32–36.
[44] M.W. Davidson, ZEISS Microscopy Online Campus | Tungsten-Halogen Lamps, Available at: http://zeiss-campus.magnet.fsu.edu/articles/lightsources/tungstenhalogen.html (accessed 15.01.16).
[45] Labsphere, Inc., Spectralon Diffuse Reflectance Standards, 2016, Available at: https://www.labsphere.com/site/assets/files/2628/spectralon_diffuse_reflectance_standards.pdf (accessed 30.11.16).
[46] eXtension.org, What's the Difference Between a Supervised and Unsupervised Image Classification? eXtension.org, 2013, 17 October, Available at: http://articles.extension.org/pages/40214/whats-the-difference-between-a-supervised-and-unsupervised-image-classification (accessed 18.03.16).
[47] C.-W. Hsu, C.-C. Chang, C.-J. Lin, A Practical Guide to Support Vector Classification, National Taiwan University, Taiwan, 2010.
[48] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, Cambridge, 2000.
[49] P. Du, G. Li, W. Zhang, X. Wang, H. Sun, Consistency measure of multiple classifiers for land cover classification by remote sensing image, in: Multiple Classifier Systems: 8th International Workshop, MCS 2009, Proceedings, Reykjavik, Iceland, June 10–12, 2009.
[50] MathWorks, Support Vector Machines (SVM), MathWorks, Available at: http://uk.mathworks.com/help/stats/support-vector-machines-svm.html (accessed 19.01.16).
[51] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines ACM Trans. Intell. Syst. Technol. 2 (3) (2011), 27:1–27:27.