



# A smartphone image-based method for the colorimetric characterization of historical pigments in mural paintings <sup>☆</sup>



Roberto Sáez-Hernández <sup>a</sup>, María Josefa Luque <sup>b</sup>, Ángel Morales-Rubio <sup>a</sup>,  
M. Luisa Cervera <sup>a</sup>, Adela R. Mauri-Aucejo <sup>a,\*</sup>

<sup>a</sup> Department of Analytical Chemistry, Faculty of Chemistry, University of Valencia. C/Dr. Moliner, 50, Burjassot, 46100, Spain

<sup>b</sup> Department of Optics and Optometry and Vision Sciences, Faculty of Chemistry, University of Valencia. C/Dr. Moliner, 50, Burjassot, 46100, Spain

## ARTICLE INFO

### Method name:

A smartphone image-based method for the colorimetric characterization of historical pigments in mural paintings

### Keywords:

Camera characterisation  
Archaeometry  
Support vector classifier  
Green earths  
Azurite  
Egyptian blue  
Lapis lazuli  
Orpiment  
Yellow ochre  
Green malachite

## ABSTRACT

In this article, the detailed description of a smartphone-based analytical method to turn a smartphone into a colorimeter is described and applied to mural painting replicas. The method consists of a camera characterization process, to estimate CIELAB values from RGB camera responses, correcting for potential experimental interferences, like the effect of the lighting/capture geometry or software for color modification and enhancement. Next, the estimated CIELAB descriptors were used to build a chemometric Support Vector Classifier to objectively distinguish among the colors of the different pigments. The method was validated in terms of color reproduction and classification performance.

- A smartphone-based method is described to capture mural painting color.
- A color gamut-specific camera characterization procedure is developed and the performance of the smartphone-based colorimeter is compared with a spectrophotometer.
- A classification method based on the estimated CIELAB descriptors is validated in terms of classification performance.

## Specifications table

Subject area:	Chemistry
More specific subject area:	Analytical Chemistry, Image Treatment, Cultural Heritage
Name of your method:	A smartphone image-based method for the colorimetric characterization of historical pigments in mural paintings
Name and reference of original method:	N/A
Resource availability:	Colorlab Matlab toolbox: J. Malo & M.J. Luque. ColorLab: the Matlab toolbox for Colorimetry and Color Vision. Univ. Valencia 2002. <a href="http://isp.uv.es/code/visioncolor/colorlab.html">http://isp.uv.es/code/visioncolor/colorlab.html</a>

<sup>☆</sup> **Related research article:** R. Sáez-Hernández, M. Luisa Cervera, Á. Morales-Rubio, M.J. Luque, I. Pérez-Torralba, G. Gallelo, K.U. Antela, A.R. Mauri-Aucejo. Digital image-based method to identify historical pigments in wall paintings, *Dyes and Pigments*, 222 (2024) 111,912. <https://doi.org/10.1016/j.dyepig.2023.111912>.

DOI of original article: [10.1016/j.dyepig.2023.111912](https://doi.org/10.1016/j.dyepig.2023.111912)

\* Corresponding author.

E-mail addresses: [roberto.saez@uv.es](mailto:roberto.saez@uv.es) (R. Sáez-Hernández), [adela.mauri@uv.es](mailto:adela.mauri@uv.es) (A.R. Mauri-Aucejo).

<https://doi.org/10.1016/j.mex.2024.102746>

Received 3 February 2024; Accepted 3 May 2024

Available online 17 May 2024

2215-0161/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license

(<http://creativecommons.org/licenses/by-nc/4.0/>)

## Method details

### Contextualization and aims

The implementation of image-based techniques using daily technology has been a growing topic during the last years [1–3], as they provide significant advantages like the wide availability and affordability. Among other devices, smartphones have gained special attention in the field of analytical chemistry, as they represent a key device in today's technology [4,5]. Their application, despite yielding fast and accessible results that can be applied in situ, requires to take into consideration some methodological aspects to ensure the quality of the processed data. As the signal depends on different experimental variables (distance, angle, illumination, device...) the development of smartphone-based methods must be carefully tuned to suit the necessities of the case study. Nevertheless, the conversion of smartphones into reliable colorimeters presents the advantages of providing a fast, cheap and non-invasive method in the investigation of artworks.

In this context, this paper describes the analytical method used by Sáez-Hernández and co-workers [6] to estimate the CIELAB values of a sample from the RGB values of an image captured with a colorimetrically characterized smartphone in order to distinguish among different historical inorganic pigments used in mural paintings using a chemometric approach.

The different stages of the procedure are described in Fig. 1. Each section of this paper describes the different procedures and parameters to be considered and tuned for the correct implementation of a particular stage of the method.

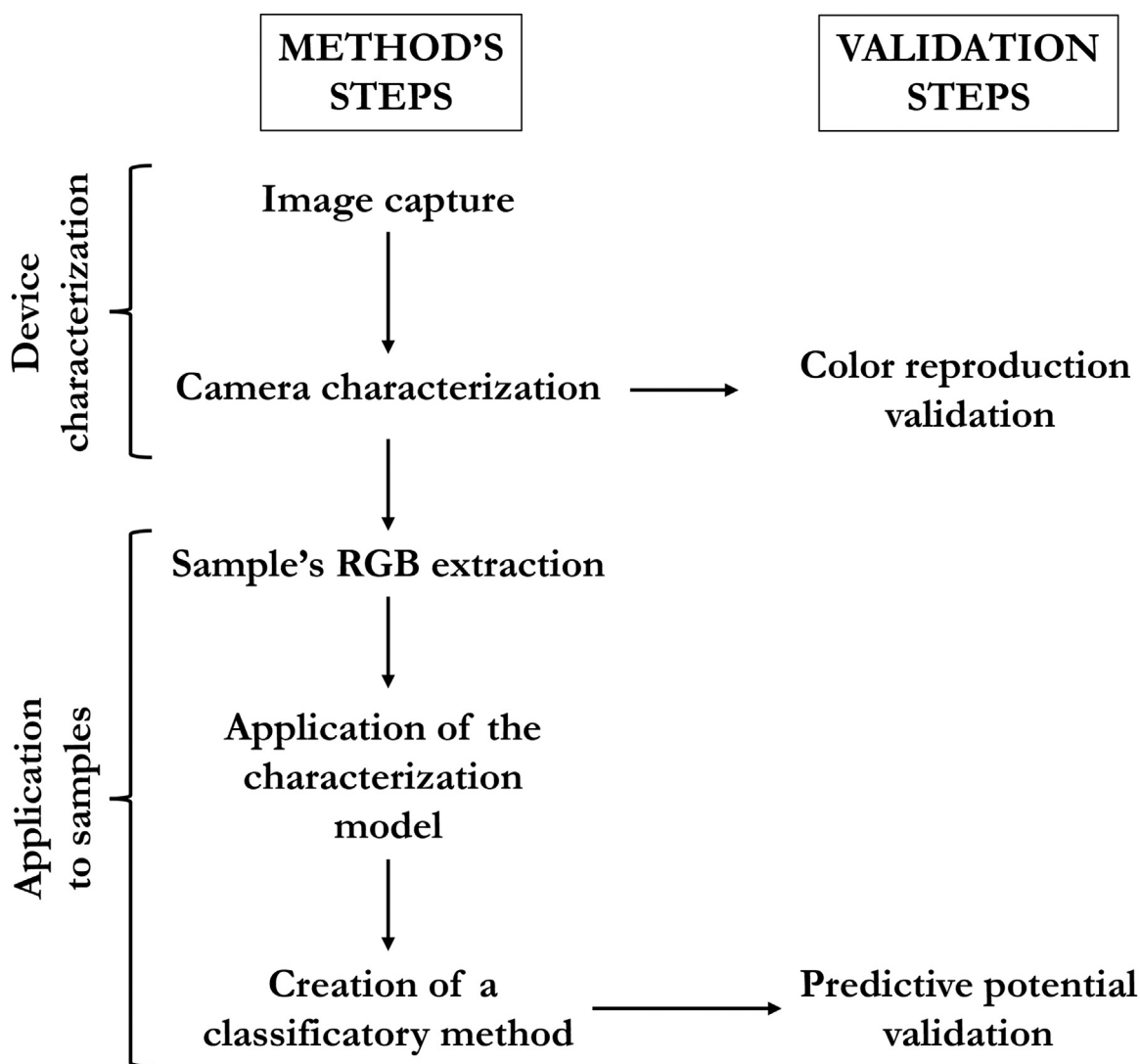
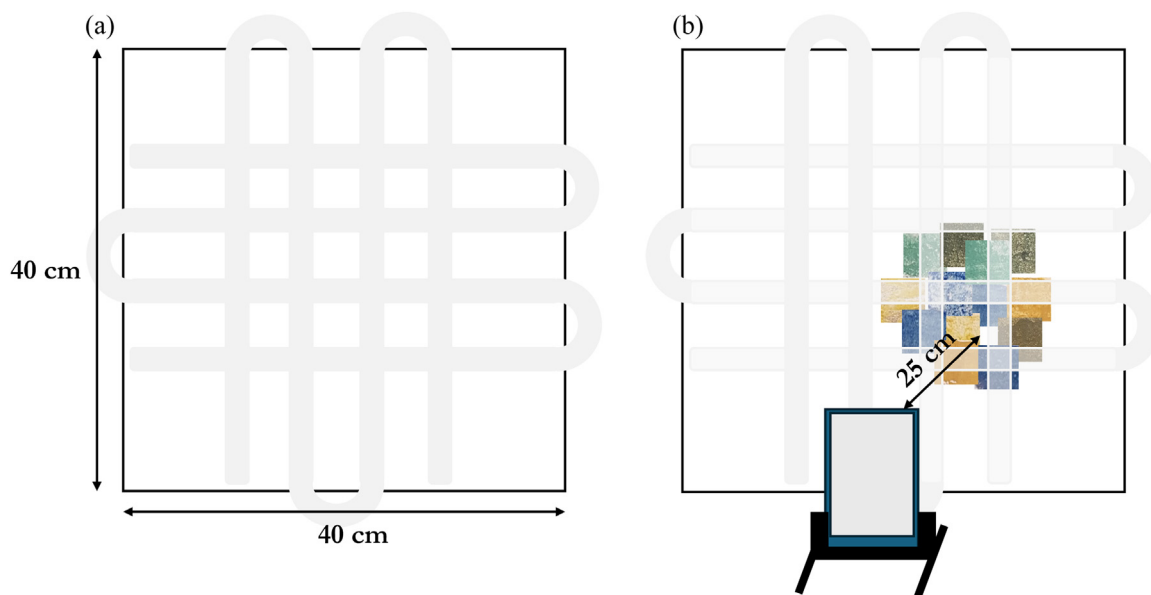


Fig. 1. Description of the overall method: first, the characterization of the camera is carried out using polynomial models; second, the color information of the data is extracted and analyzed using these models. On the right, the validation steps to assess the model's quality.



**Fig. 2.** a) Illumination board: in grey, the white LED strip arranged in a transparent 40×40 cm flat surface. b) Relative position of the smartphone, the illumination board and the samples.

#### *Image capture – setup and smartphone conditions*

The first step of the method requires to take an image of the sample under investigation. For the case reported in [6], the built-in camera application in the “pro” mode of Samsung Galaxy Edge S7 smartphone (model SM-G93F with a 12.2 MP camera sensor) was used. ISO was set to 100, aperture to 1/180 and white balance was 4000 K. The images were captured in .jpeg format. However, this procedure can be carried out using any device that allows to control camera settings, to choose those that provide greater similarity between the observed image and the real scene. The smartphone was placed on a lab-made 3D-printed support that allows the smartphone to be held parallel to the sample.

Samples were illuminated with a white LED strip arranged on a flat surface placed at a distance of 25 cm from the samples and placing the smartphone parallel to the painted surface in a darkened environment. This source of illumination is recommended to help controlling lighting environment. In this case, the strip was stuck to a 40 x 40 cm transparent surface creating a mesh, as observed in Fig. 2. This illumination board was held perpendicular to the horizontal using a couple supports on each side, ensuring homogenous illumination on the surface of the sample. The smartphone, placed on the support, was located at the same plane as the illumination board, as depicted in Fig. 2. Using the proposed procedure, a surface of 20 x 20 cm of painted area can be analyzed, although this area will depend upon the properties of each camera.

The color characterization model is trained from color data of a color chart, captured under the same illumination conditions as the problem samples. This color chart consists of a set of different color chips covering the region of color space of the problem samples. In this case, a commercial color chart is used (ColorChecker Classic, XRite 24 colors, 57 × 86 mm) containing a set of 18 colored chips and a grey scale ranging from white to black (Fig. 3). Using a spectrophotometer (Konica Minolta Sensing, CM-26d; D65 reference illuminant, 10° standard observer, in the Specular Component Excluded (SCE) mode), the CIELAB descriptors of each point of the chart are measured (Table 1). The chart’s RGB and CIELAB data are used to create the characterization model.

This color chart must be stuck to the wall, totally parallel to the sample and the camera, and at the same plane that the painted surface, so that the distance between the camera and the samples equals the distance from the camera to the color chart.

#### *Camera’s colorimetric characterization model*

The RGB values from a circular region are obtained from each sample in the color chart, using a set of Matlab routines. These are based on the following steps:

- (i) Manually select the region to analyze
- (ii) Upon clicking, the software selects a 50-pixel radius circle around that point
- (iii) Average the RGB values of the pixels inside that circle.

The procedure is repeated for each of the 24 color chips, obtaining the mean of the RGB values of these chips. A polynomial model is created correlating the real CIELAB descriptors (seen in Table 1) with the RGB values as obtained for each colored spot. The best



Fig. 3. ColorChart employed for this method. Numbers indicate the ordinal assignment to each color, as corresponds to Table 1.

Table 1  
 CIELAB descriptors of the color chips of the ColorChecker chart. Sample are labelled as in Fig. 3.

Color	$L^*$	$a^*$	$b^*$	Color	$L^*$	$a^*$	$b^*$
1	70,0	-29,0	4,2	13	52,4	-27,8	-20,3
2	55,8	6,1	-19,9	14	51,7	39,8	-13,8
3	43,7	-10,0	19,6	15	77,5	5,0	74,0
4	50,4	-3,9	-18,5	16	41,7	41,2	21,3
5	63,3	13,0	17,3	17	54,9	-32,9	31,6
6	39,4	11,0	11,8	18	33,6	9,7	-39,3
7	67,7	19,1	59,8	19	25,0	-0,1	-0,6
8	69,5	-19,2	54,8	20	37,9	-0,5	0,0
9	31,8	16,6	-17,6	21	51,1	0,1	1,1
10	49,5	41,3	14,0	22	66,2	-0,5	1,5
11	42,6	6,9	-37,7	23	79,6	-0,7	2,7
12	59,2	33,0	49,8	24	92,8	-0,8	4,8

fit is obtained with a 2-degree polynomial using the pseudoinverse procedure in Eq. (1).

$$H = P_n(R, G, B) \cdot M_n \tag{1}$$

where  $H$  is a  $N \times 3$  matrix (in this case,  $N = 24$ ) of experimental  $L^*a^*b^*$  values, each row containing a color in the chart and each column a color descriptor,  $L^*$ ,  $a^*$  and  $b^*$ . Next,  $P_n(R, G, B)$  is a  $N \times K$  matrix, containing the different terms of a polynomial of degree “ $n$ ” in  $R$ ,  $G$  and  $B$  (for instance, if  $n = 2$ , the  $K = 10$  columns contain the terms 1,  $R$ ,  $G$ ,  $B$ ,  $RG$ ,  $RB$ ,  $GB$ ,  $R^2$ ,  $G^2$  and  $B^2$ , respectively) and  $M_n$  represents the  $K \times 3$  matrix of coefficients of the polynomial.

The degree of the polynomial was selected based on two main factors. First, the predictive capacity of the model to yield correct color descriptors, and second, the potential overfitting that might arise when using unnecessarily complex models. Check section **Validation of the characterization of the camera** for the results of this selection process.

The camera’s characterization model intends to account for the sensitivities of the RGB sensors, the influence of the lighting conditions and the effects of the image enhancement software of the camera, so that real CIELAB descriptors can be obtained from the image. Therefore, to increase the robustness of the procedure, matrix  $M_n$  (see equation [1]) is computed for each photograph. In

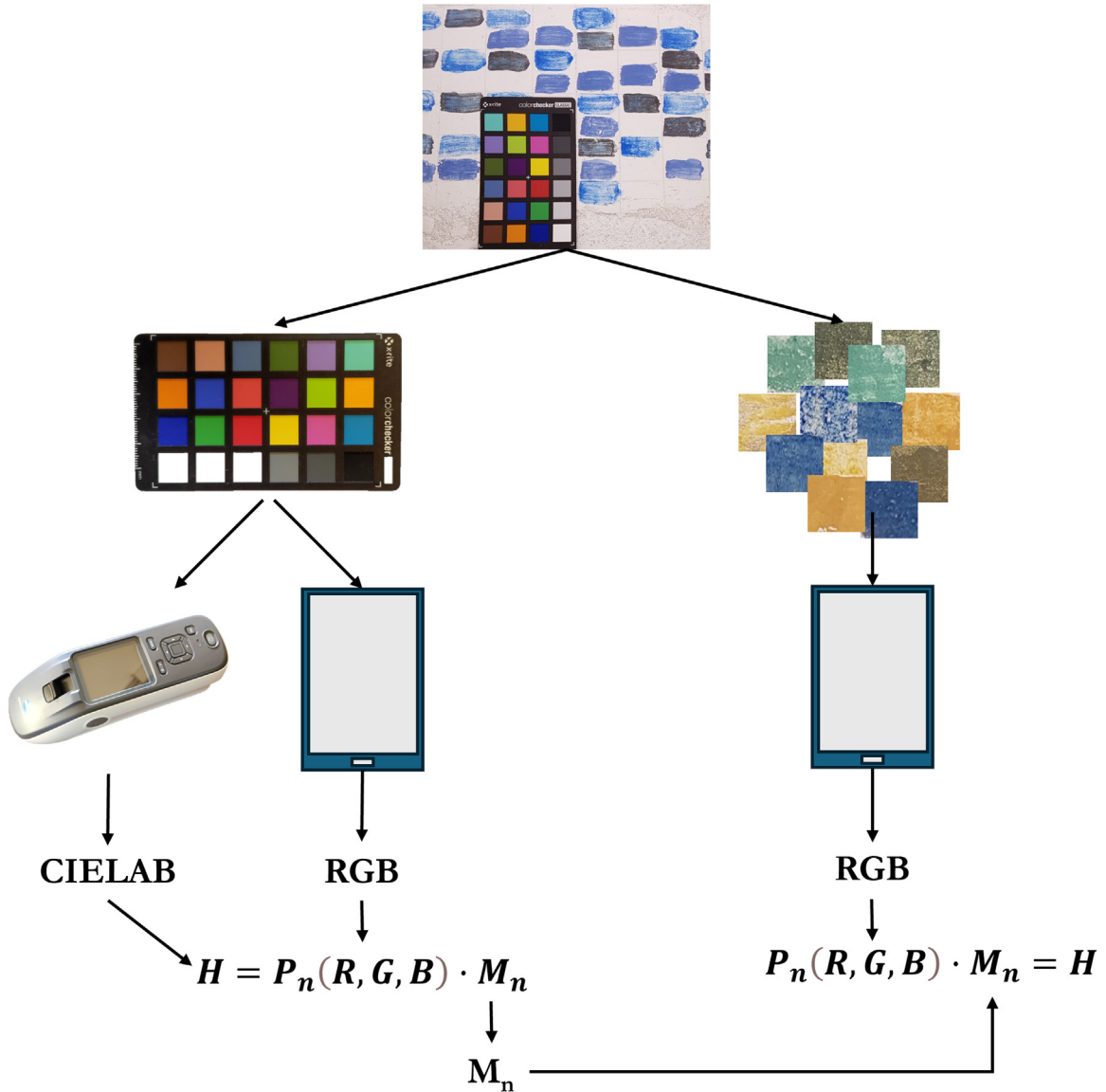


Fig. 4. Schematic representation of the camera characterization and signal treatment to obtain corrected CIELAB descriptors of the painted spots.

this way, slight changes in illumination among photographs of the same scene or the influence of a change in the objects included in the scene will be accounted for in the model obtained from that specific photograph. At a practical level, this means that the coefficients,  $M_n$ , are idiosyncratic of each one of the photographs. In Fig. 4 there is a schematic representation of the camera characterization and signal treatment as described in this section.

#### Samples' RGB extraction and application of the characterization model

Once the camera has been characterized and the correction model has been created as described above, the painted surfaces can be studied from their colorimetric properties perspective. To that end, different spots are segmented following the same approach explained before, ensuring that the sampling represents the colored spot and not sampling the same area more than once. These RGB values are now inputted in Eq. (1), using the polynomial of degree 2 ( $P_2$ ) and the coefficients ( $M_2$ ) obtained for their respective photographs (as these represent the experimental conditions that correct the significant variables like lighting environment or color enhancement). As an output of the computation, a set of corrected CIELAB descriptors will be obtained (matrix  $H$ , which will contain as many rows as samples, and 3 columns:  $L^*$ ,  $a^*$  and  $b^*$ ).

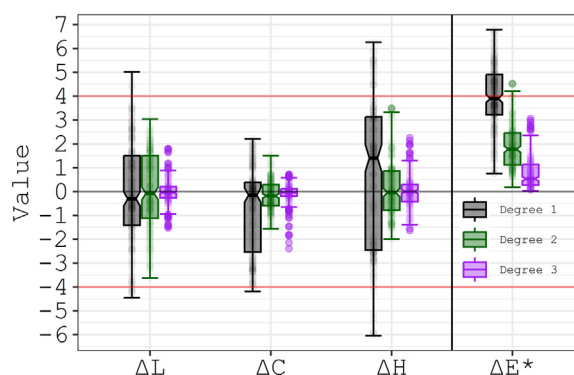


Fig. 5. CIE $\Delta E_{2000}$  values obtained for the comparison of the three different polynomials tested.

Table 2

Pairwise comparison of the color differences obtained with the 3 different models. If the confidence interval of the mean difference includes 0, both groups do not show statistically significant differences; otherwise, both groups are considered as having significant difference.

Degrees compared	Difference between the estimated group means	Confidence interval for the true mean difference (95%)	<i>p</i> -value
1 – 2	2.75	[1.73 – 3.76]	$6.7 \cdot 10^{-10}$
1 – 3	5.19	[4.18 – 6.20]	0
2 – 3	2.45	[1.43 – 3.46]	$4.8 \cdot 10^{-8}$

### Creation of a classification method

Using the corrected CIELAB values obtained as described in section “**Samples’ RGB extraction and application of the characterization model**”, a Support Vector Classifier (SVC) model is created. In this work, the method has been developed for three different color hues: blue, green and yellow. Each hue was represented by three different inorganic species. Thus, one SVC model was created for each hue, obtaining a total of three SVC. The reason for this choice is that this method aims to help in the distinction among similar colors painted with different pigments: hence, it is necessary, for instance, to distinguish different shades of blue but not between blue and yellow.

SVC models need to be optimized in advance for different parameters and variables. In this article, all three SVC used the radial kernel (as it is needed to expand the classification boundaries further from two categories: in this case, three). This kernel uses an intrinsic value ( $\gamma$ ) that also requires optimization, and for this case study, their optimal values were 0.5, 3 and 0.5 for yellow, blue and green, respectively (see the reference article for the discussion about the influence of each variable on predictive capacity [6]). To create these models, *R* [7] was used, and *pROC* [8], *verification* [9] and *e1071* [10] packages were employed. Graphical representation was carried out using *ggplot2* [11].

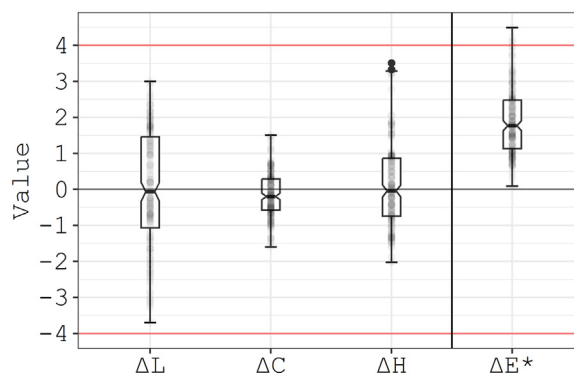
### Method validation results

#### Selection of the degree of the polynomial

Three models were tested, with 1st, 2nd and 3rd degree polynomials respectively using data from 10 captures of the reference color chart. The color differences between the predictions of the models and the experimental  $L^*a^*b^*$  values for each sample were computed using the CIE $\Delta E_{2000}$  color difference formula [12]. The results are shown in Fig. 5. Besides the global color difference,  $\Delta E^*$ , Fig. 5 includes the differences in chroma ( $\Delta C^*$ ), in lightness ( $\Delta L^*$ ) and in hue ( $\Delta H^*$ ), as described in the work of Luo et al. [12].

Acceptable color differences must be below 4.0 CIE $\Delta E_{2000}$  units, which is the tolerance limit proposed for printed materials [13]. The Friedman test indicates significant differences in  $\Delta E^*$  between the models (chi-square 431, *p*-value  $\approx 0$ ). A Tukey-Kramer post-hoc test (Table 2), shows that all pairwise comparisons are statistically significant, with significantly lower  $\Delta E^*$  values when the degree of the polynomial is  $n = 3$  and worse results for  $n = 1$ , as suggested in Fig. 5.

As expected, increasing the degree,  $n$ , in the polynomial reduces the dispersion of the result and the magnitude of the color errors, since the median value of  $\Delta E^*$  gets closer to 0 in Fig. 5. Acceptable  $\Delta E^*$  results within the tolerance limit of 4 units are obtained for  $n \geq 2$ . Nonetheless, the 2nd degree polynomial was selected following the parsimony principle. This idea states that when selecting between two potential methods that yield good results, the simpler one is preferred, as it avoids potential overfitting. Overfitting accounts for the creation of very complex models that get tuned exclusively to the dataset used for its creation, meaning that the interpolation of new samples not originally considered could lead to high prediction errors. In this paper, as both 2nd and 3rd degree



**Fig. 6.** Distribution of  $\Delta L^*$ ,  $\Delta C^*$ ,  $\Delta H^*$  and  $\Delta E^*$  computed values for the validation of the camera characterization procedure ( $n = 240$ ). Differences were calculated as (value provided by the reference device - value provided by the proposed method). In red, the threshold value as established by the literature [13].

**Table 3**  
Composition of the dataset.

Hue	Pigment	Total number of samples	Number of samples per hue	Training set	Validation set	Prediction error /%
Green	Lead-tin yellow	80	213	170	43	2.9
	Yellow ochre	68				
	Orpiment	65				
Blue	Egyptian blue	90	272	218	54	2.2
	Azurite	90				
	Lapis lazuli	92				
Yellow	Celadonite	65	203	162	41	5.7
	Green malachite	70				
	Verona green	68				

polynomials provided good analytical results in terms of color reproduction, the 2nd degree was selected. Note that in each image, data from 24 color samples are used to derive 10 parameters when  $n = 2$  but 20 when  $n = 3$ .

An inspection of the outliers allowed to observe that degree 1 presented no outliers (defined as any observation out of the range comprised between the 1st quartile minus 1.5 times the interquartile range and the 3rd quartile plus 1.5 times the interquartile range), degree 2 presented 1 outlier (0.4% in total) on the upper end, and degree 3 had 10 outliers (4.2% in total) on the upper end. This investigation also advises against the use of the 3rd degree, as it yields more outliers.

#### Validation of the characterization of the camera

To validate the performance of the smartphone as a colorimeter under the working conditions, the computed CIELAB data of the color chart spots are compared to their real CIELAB values (shown in Table 1). To that end, the CIE $\Delta E$ 2000 values shown in Fig. 5 for the model of 2nd degree are depicted in Fig. 6 for clarity.

Fig. 6 shows that  $\Delta E^*$ , with a median of  $1.8 \pm 1.4$  units, is within the tolerance limits. This means that, overall, the method described here is able to acceptably measure color from smartphone images when compared to a reference spectrophotometer. The lightness differences  $\Delta L^*$  have a centered distribution very close to 0 (median value  $-0.1$ , Interquartile Range (IQR) of 2.5), meaning that the corrected lightness ( $L^*$ ) values reproduced by the smartphone are comparable to the ones measured by the colorimeter. Chroma differences,  $\Delta C^*$ , have a median of  $-0.2$  CIE $\Delta E$ 2000 units, with an IQR of 0.9. Lastly,  $\Delta H^*$ 's median value was found to be  $0.0 \pm 1.6$  CIE $\Delta E$ 2000 units. The results shown for  $\Delta C^*$  and  $\Delta H^*$  demonstrate that the method colorimetric errors are within the accepted tolerance, with larger dispersion in lightness  $L^*$  than in chroma and hue ( $C^*$  and  $H^*$ ). Also, the fact that all three differences are found centered at 0 means that the proposed method is not biased towards a specific direction, that is, it does not systematically overestimates or underestimates neither of the parameters.

#### Validation of the predictive capacity of the chemometric model

Using the information obtained by the image-based method, three different classificatory methods were built and validated. To that end, the dataset (shown in Table 3) was split into an 80 – 20 proportion to be used for the model training and validation, respectively.

According to the validation data shown in Table 3, the performance of the classification method based on color parameters is confirmed. For all three hues, the models demonstrated a great predictive potential allowing to objectively distinguish among the different pigments.

### Ethics statements

Non-applicable.

### Declaration of Competing Interest

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

### CRedit authorship contribution statement

**Roberto Sáez-Hernández:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **María Josefa Luque:** Software, Validation, Writing – review & editing, Supervision, Funding acquisition. **Ángel Morales-Rubio:** Investigation, Resources, Supervision, Writing – review & editing. **M. Luisa Cervera:** Investigation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Adela R. Mauri-Aucejo:** Investigation, Writing – review & editing, Supervision, Funding acquisition.

### Data availability

Data will be made available on request.

### Acknowledgments

Funding: This work was supported by the Valencian Government – “Conselleria d’Innovació, Universitats, Ciència i Societat Digital” [PROMETEO-2019–056] and the Spanish Government – Ministry of Universities [FPU19/02304].

### References

- [1] K. Grudpan, S.D. Kolev, S. Lapanantnopakhun, I.D. McKelvie, W. Wongwilai, Applications of everyday IT and communications devices in modern analytical chemistry: a review, *Talanta* 136 (2015) 84–94, doi:10.1016/j.talanta.2014.12.042.
- [2] G.M. Fernandes, W.R. Silva, D.N. Barreto, R.S. Lamarca, P.C.F. Lima Gomes, J. Flávio da S Petrucci, A.D. Batista, Novel approaches for colorimetric measurements in analytical chemistry – a review, *Anal. Chim. Acta* 1135 (2020) 187–203, doi:10.1016/j.aca.2020.07.030.
- [3] L.F. Capitán-Vallvey, N. López-Ruiz, A. Martínez-Olmos, M.M. Erenas, A.J. Palma, Recent developments in computer vision-based analytical chemistry: a tutorial review, *Anal. Chim. Acta* 899 (2015) 23–56, doi:10.1016/j.aca.2015.10.009.
- [4] P. Yadav, L. Yadav, H. Laddha, M. Agarwal, R. Gupta, Upsurgence of smartphone as an economical, portable, and consumer-friendly analytical device/interface platform for digital sensing of hazardous environmental ions, *Trend. Environ. Anal. Chem.* 36 (2022) e00177, doi:10.1016/j.teac.2022.e00177.
- [5] S. Soares, G.M. Fernandes, F.R.P. Rocha, Smartphone-based digital images in analytical chemistry: why, when, and how to use, *TrAC Trend. Anal. Chem.* 168 (2023) 117284, doi:10.1016/j.trac.2023.117284.
- [6] R. Sáez-Hernández, M.L. Cervera, Á. Morales-Rubio, M.J. Luque, I. Pérez-Torralba, G. Gallelo, K.U. Antela, A.R. Mauri-Aucejo, Digital image-based method to identify historical pigments in wall paintings, *Dyes Pigment.* 222 (2024) 111912, doi:10.1016/j.dyepig.2023.111912.
- [7] R. Core-Team, R: a language and environment for statistical computing, (2022). <https://www.r-project.org/>.
- [8] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, *BMC Bioinform.* 12 (2011) 77, doi:10.1186/1471-2105-12-77.
- [9] r NCAR - Research Applications Laboratory, verification: Weather Forecast Verification Utilities, (2015). <https://CRAN.R-project.org/package=verification>.
- [10] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, F. Leisch, e1071: misc functions of the department of statistics, Probab. Theory Group (Formerly: E1071) (2021) <https://cran.r-project.org/package=e1071>.
- [11] H. Wickham, ggplot2: elegant graphics for data analysis, (2016).
- [12] M.R. Luo, G. Cui, B. Rigg, The development of the CIE 2000 colour-difference formula: CIEDE2000, *Color Res. Appl.* 26 (2001) 340–350. <https://doi.org/10.1002/col.1049>.
- [13] H.X. Liu, B. Wu, Y. Liu, M. Huang, Y.F. Xu, A discussion on printing color difference tolerance by CIEDE2000 color difference formula, *Appl. Mech. Mater.* 262 (2013) 96–99, doi:10.4028/www.scientific.net/AMM.262.96.