Molada Tebar, A.; Lerma García, JL.; Marqués Mateu, Á. (2018). Camera characterization for improving color archaeological documentation. Color Research and Application. 43(1):47-57. doi:10.1002/col.22152

CAMERA CHARACTERISATION FOR IMPROVING COLOUR ARCHAEOLOGICAL DOCUMENTATION

Molada-Tebar, Adolfo* Lerma, José Luis Marqués-Mateu, Ángel

Department of Cartographic Engineering, Geodesy, and Photogrammetry, Universitat Politècnica de València. 46022 Valencia, Spain – admote@doctor.upv.es

ABSTRACT:

Determining the correct colour is essential for proper cultural heritage documentation and cataloguing. However, the methodology used in most cases limits the results since it is based either on perceptual procedures or on the application of colour profiles in digital processing software. The objective of this study is to establish a rigorous procedure, from the colourimetric point of view, for the characterisation of cameras, following different polynomial models. Once the camera is characterised, users obtain output images in the sRGB space that is independent of the sensor of the camera. In this paper we report on pyColourimetry software that was developed and tested taking into account the recommendations of the *Commission Internationale de l'Éclairage* (CIE). This software allows users to control the entire digital image processing and the colourimetric data workflow, including the rigorous processing of raw data. We applied the methodology on a picture targeting Levantine rock art motifs in Remigia Cave (Spain) that is considered part of a UNESCO World Heritage Site. Three polynomial models were tested for the transformation between colour spaces. The outcomes obtained were satisfactory and promising, especially with RAW files. The best results were obtained with a second order polynomial model, achieving residuals below three CIELAB units. We highlight several factors that must be taken into account, such as the geometry of the shot and the light conditions, which are determining factors for the correct characterisation of a digital camera.

KEYWORDS: rock art, documentation, colourimetry, CIE colour spaces, Python

1. INTRODUCTION

Colour is a fundamental feature for proper cultural heritage documentation in general, and archaeological rock art documentation in particular. The correct determination of colour provides vital information, not only at a descriptive level, but also, especially, at a technical and quantitative level. This allows a better understanding of the study area and provides useful information regarding the origin and aging of the pigments.

Traditionally, rock art documentation methods were restricted to subjective procedures based on direct observations of the researcher with the use of colour charts (for instance the Munsell Color Book). Despite its clear advantages, this methodology entails practical and technical limitations, affecting the results obtained in determining the colour [1].

The problem is that colour is a matter of perception and subjective interpretation. There is no single physical and universal scale for its measurement [2,3]. In the particular case of visual assessments, even if several observers examine the same object, they will obtain different references and experiences and may express the same colour stimulus with completely different words. It is obvious that verbal expressions cannot be used to communicate colour information.

In recent times, it is becoming more frequent to combine classical heritage documentation techniques based on perceptual procedures with rigorous procedures, supported by digital images [4,5]. In this case, it is necessary to use rigorous colourimetric techniques to properly support rock art research. In these applications, measurements must be carried out based on direct measurements using colourimeters or spectrophotometers. Moreover, the rigorous processing of colourimetric data requires software packages with colourimetric specific technical characteristics.

There are different software options, both for the treatment of colourimetric samples, generally provided by the manufacturer of the instrument used in the measurement, as well as for the digital image processing [6]. Although in both cases they are, to some extent, versatile software systems with wide functionality, they generally do not allow the user to have absolute control on the methodological process. Actually, it is not usual to find software that allows raw processing RGB data, characterising digital cameras, and specially applying characterisation parameters on input images to obtain characterised output sRGB images.

Given the importance of colour communication in archaeology, we have developed our own software package for the treatment of colourimetric and spectral data, named pyColourimetry. This software allows users to apply a procedure for the characterisation of digital cameras to collect quantitative colourimetric information in physically based colour spaces [7-9]. The idea is to achieve colour information independently of the device used in the data acquisition stage, relying on digital images and colourimetric measurements only.

By characterisation we refer to the determination of the transformation equations so that the acquired RGB information is brought into physically based colour spaces such as those introduced by the *Commission Internationale de l'Éclairage* (CIE), that is, the CIE XYZ space and its derivatives. In this way, a conventional digital camera could be used for rigorous colour determination, somehow simulating a colourimeter [10-12]. Once characterised, we can work in a physical colour space, which is independent of the device and comparable with other devices already characterised.

Different methods, including polynomial [12], principal component analysis [13] or artificial neural networks [14], can be used to characterise digital image devices. In our software we implemented the polynomial method because it gives substantially the same results as the other methods at lower pre-processing and computational cost. In particular, to obtain the RGB-CIE XYZ transformation equations, three polynomial models were considered: linear, second and third order.

The proposed method is based on objective methods that are independent of the observer experience. It combines the direct method, based on colourimetric measurements obtained using specific instruments (colourimeter and spectrophotometer), and the indirect method, using digital images with RGB information. Results from previous research show that good results can be obtained using polynomial transformation equations, especially those of second or third order [11,12]. It is not a restrictive methodology, and can be used in a complementary way to other techniques for the study and conservation of rock art specimens, for instance, laser scanning together with photogrammetric techniques, which allow the generation of 3D photorealistic models of rock art paintings [15].

Our basic idea is to adapt pyColourimetry functionalities to specific needs of both information users and information providers. pyColourimetry software also allows us to achieve a full control of the established methodological process.

2. CIE COLOUR SPACES

Since 1931, the CIE has developed systems to express colour numerically. Colour spaces define the quantitative relationship between physical pure colours in the electromagnetic visible spectrum, and physiological perceived colours in human vision. The mathematical relationships that define these colour spaces are essential tools for colour management [16].

Two well-known colour spaces that provide consistent approaches in relation to the human visual system are the CIE xyY and the CIE L*a*b*, both based on the so-called CIE XYZ tristimulus values.

2.1 CIE XYZ

The CIE 1931 XYZ colour space allows any colour stimulus (usually expressed in terms of radiance at fixed wavelength intervals in the visible spectrum) to be represented with three parameters XYZ called tristimulus values. CIE XYZ tristimulus values are fundamental measures of colour and are directly used in a number of colour management operations. This colour space serves as a reference to define many other colour spaces.

It is based on the additive colour mixing principle. All colour signals can be matched by the additive mixture of three primaries. In this colour space, the primary colours used are not real colours, in the sense that they cannot be generated with any light spectrum.

The second coordinate Y represents the luminance, that is the total radiation reflected in the visible spectrum. Z is quasi-equal to blue stimulation (or the S cone response of the human eye), and X is a linear combination of cone response curves chosen to be nonnegative [12,16].

The CIE XYZ tristimulus values (Eq.1) can be obtained as follows [16]:

$$\begin{aligned} X &= k \sum_{\lambda} \phi_{\lambda}(\lambda)\bar{x}(\lambda)\Delta\lambda \\ Y &= k \textstyle\sum_{\lambda} \phi_{\lambda}(\lambda)\bar{y}(\lambda)\Delta\lambda \\ Z &= k \sum_{\lambda} \phi_{\lambda}(\lambda)\bar{z}(\lambda)\Delta\lambda \end{aligned} \qquad (1)$$

$$k = 100/\textstyle\sum_{\lambda} S(\lambda)\bar{y}(\lambda)\Delta\lambda$$

where $\phi_{\lambda}(\lambda)$ denotes the spectral distribution of the colour stimulus function; $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ are colour-matching functions of a standard colourimetric observer; k is a normalising constant; and $S(\lambda)$ is the relative spectral power distribution of the illuminant.

2.2 CIELAB

The three-dimensional colour space produced by plotting CIE tristimulus values XYZ in

rectangular coordinates is not visually uniform [16]. Equal distances in this space do not represent equally perceptible differences between colour stimuli. For this reason, in 1976, the CIE introduced two new colour spaces (CIELAB and CIELUV) whose coordinates are non-linear functions of X, Y and Z. This non-linear transform of the XYZ values provided partial solutions to both the problems of colour appearance and colour difference.

The CIE 1976 L*a*b* colour space provides a three-dimensional colour space where the a*-b* axes form a plane to which the L* axis is orthogonal. It separates the colour information into lightness (L∗) and colour information (a*, b*) on two a red/green (a*) and yellow/blue (b*) axes. The lightness of a colour stimulus ranges from 0 representing black to 100 representing white. As the position of a colour moves from the central region toward the edge of the sphere, its saturation (or chroma) increases.

The transformation from tristimulus values to L*a*b* coordinates is given by the well-known equation (Eq. 2) [16]:

$$\begin{aligned} L^* &= 116 \cdot f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* &= 500 \cdot \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \\ b^* &= 200 \cdot \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \end{aligned} \qquad (2)$$

where

$$f\left(\frac{X}{X_n}\right) = \begin{cases} \left(\frac{X}{X_n}\right)^{\frac{1}{3}}, & \text{if } \left(\frac{X}{X_n}\right) > \left(\frac{6}{29}\right)^3 \\ \left(\frac{841}{108}\right) \cdot \left(\frac{X}{X_n}\right) + \left(\frac{4}{29}\right), & \text{if } \left(\frac{X}{X_n}\right) \leq \left(\frac{6}{29}\right)^3 \end{cases}$$

$$f\left(\frac{Y}{Y_n}\right) = \begin{cases} \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}, & \text{if } \left(\frac{Y}{Y_n}\right) > \left(\frac{6}{29}\right)^3 \\ \left(\frac{841}{108}\right) \cdot \left(\frac{Y}{Y_n}\right) + \left(\frac{4}{29}\right), & \text{if } \left(\frac{Y}{Y_n}\right) \leq \left(\frac{6}{29}\right)^3 \end{cases}$$

$$f\left(\frac{Z}{Z_n}\right) = \begin{cases} \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}}, & \text{if } \left(\frac{Z}{Z_n}\right) > \left(\frac{6}{29}\right)^3 \\ \left(\frac{841}{108}\right) \cdot \left(\frac{Z}{Z_n}\right) + \left(\frac{4}{29}\right), & \text{if } \left(\frac{Z}{Z_n}\right) \leq \left(\frac{6}{29}\right)^3 \end{cases}$$

X, Y, Z are the tristimulus values of test colour stimulus; and X_n, Y_n, Z_n are the corresponding tristimulus values of a specified white colour stimulus

2.3 *Colour differences ΔE*

The CIE XYZ colour space does not match well to the human perception of colour differences. The colour differences perceived equally in different colour regions could have very different distances in the CIE XYZ colour space [17]. In classical colourimetry, colour difference metrics are determined using formulas based on the CIELAB colour space that is more perceptually uniform than the CIE XYZ space. The non-linear

transform of tristimulus values in the CIELAB equations allows the Euclidean distance between two points in the new space to better predict the visual colour difference between the colour stimuli represented by two points (ΔE).

Given a pair of colour stimuli in CIELAB space, we establish the ΔE^*_{ab} CIELAB colour difference by the follow equation (CIE76, Eq. 3):

$$\Delta E^*_{ab} = \sqrt{(\Delta L^*_{ab})^2 + (\Delta a^*_{ab})^2 + (\Delta b^*_{ab})^2} \quad (3)$$

It is advisable to work with CIELAB differences since we work in a uniform space, whereas this condition is not fulfilled in the CIE XYZ colour space. It is for this reason that CIELAB colour differences are useful to establish colour tolerances between sample measurements and standards, and determine if the samples are considered acceptable. The criterion that allows us to establish colour tolerances is based on the concept of "just noticeable difference" (JND), that is, a hardly perceptible difference between two sensory stimuli.

If ΔE^*_{ab} approximates to 2.3, being in any case inferior to 3 [18,19] we would be considering differences of colour hardly perceptible by the human eye. Further, in cultural heritage digitalisation, we can use two guidelines that provide measures to assess the colour quality of digital images: FADGI (Federal Agencies Digitization Guidelines Initiative) from the US and Metamorfoze (Netherland's national programme for the preservation of paper heritage [22,23]) from the Netherlands. The Metamorfoze guideline measures colour quality using the CIE76 colour difference formula, and defines a tolerance average ("colour accuracy") ΔE^*_{ab} of 4 [23,24]. In this study we set the maximum acceptable value ΔE^*_{ab} to three units.

There are a number of proposed improvements to the original ΔE calculation. Although CIEDE2000 is the most recent standard to compute colour differences [25] we used the CIE76 because it fits better to typical working environments found in archaeological sites. Moreover, the CIEDE2000 report establishes a number of 'reference conditions' such as sample size (4 degrees), sample-sensor separation (contact) and sample homogeneity (textureless) which are too demanding in practical applications. None of these conditions can be totally guaranteed in archaeological sites or even in rock art specimens under laboratory conditions. Therefore, in this study we used the CIE76 colour difference formula [16] for analysing the results achieved after camera characterisation.

3. MATERIAL AND METHODS

3.1 Image data

Trichromatic digital cameras capture colour information in the well-known RGB format. The signal generated by the digital camera is device dependent. By means of the characterisation we establish the relationship between device dependent RGB values and the tristimulus coordinates defined by the CIE standard colourimetric observer.

The steps required for the characterisation of the camera are: (1) to collect an image of a colour chart; (2) to extract the RGB data from the image; (3) to obtain the CIE XYZ coordinates of the patches present in the colour chart; and (4) to compute the

transformation from RGB to CIE space. Although it is recommended to use RAW image files rather than processed or compressed image files such as JPEG or ECW, this functionality is rarely found in available packages.

The data used herein consist of a number of digital images taken on a rock art site, called *Cova Remigia* (Remigia Cave, Ares del Maestrat, Castellón). This site is one of the most singular rock art caves of the Mediterranean Basin on the Iberian Peninsula. Its vividly and graphically narrative scenes make it a unique space in Levantine rock art that was included in the UNESCO World Heritage list in 1998 [26].

A Fujifilm IS PRO camera (with a 60 mm aspherical lens) was used to take RAW images. This device has a high sensitivity and a low noise CCD sensor with a depth of 14 bits/pixel. The X-rite ColorChecker SG Digital Colour Chart with 140 patches was included in this rock art scene (Figure 1).

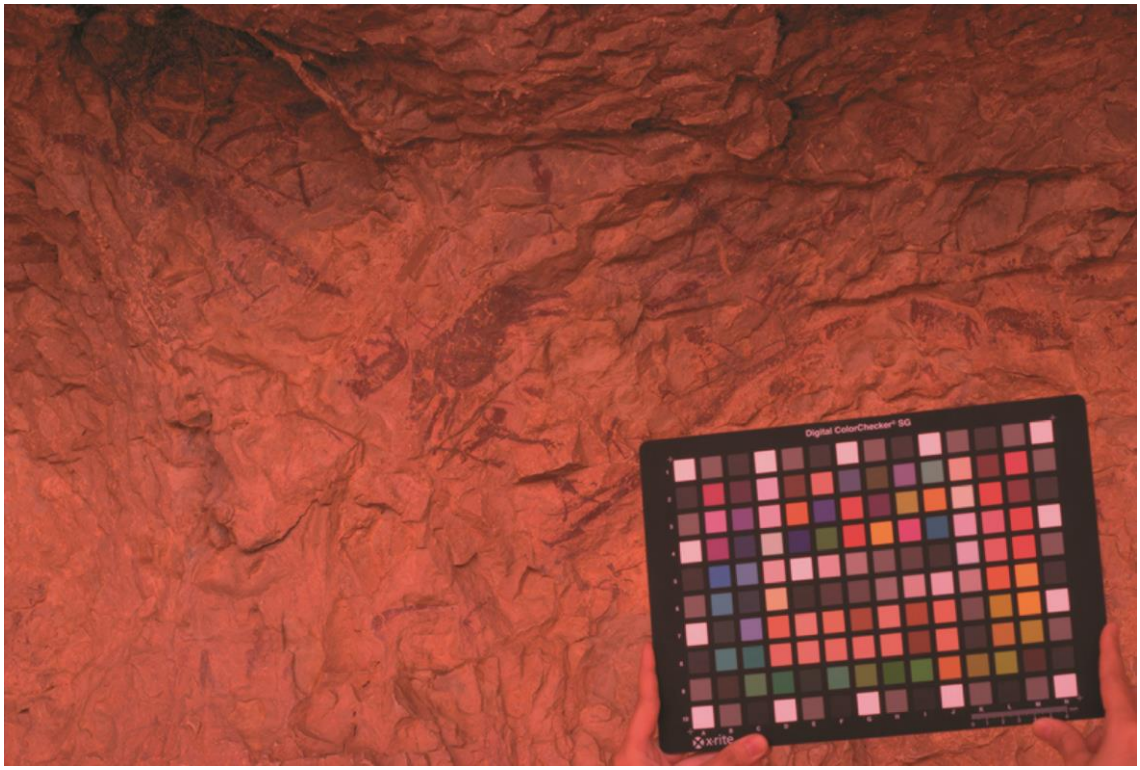


FIGURE 1 Example of data acquisition on site. Partial view of the Shelter V, Cova Remigia.

The CIE XYZ tristimulus coordinates of the colour patches are usually provided by the manufacturer. However, it is recommended to perform a direct measurement which allows us to control the conditions of the data collection, which in general differ from those of the published chart.

The instrument used in our experiments was the colorimeter Konica Minolta CS-100A, which was calibrated before the measurement sessions. An average of four measurements was obtained for each colour patch using a standard 2º observer and D65 illuminant. The data with the CIE XYZ coordinates were stored in a text file for further processing.

Finally, we obtained three subsets of samples (grey samples, training samples, and testing samples) with RGB data from a number of selected patches using the pyColourimetry

software (Figure 2). We used these RGB data sets to determine the transformation equations, as well as to evaluate the results obtained after the characterisation process.

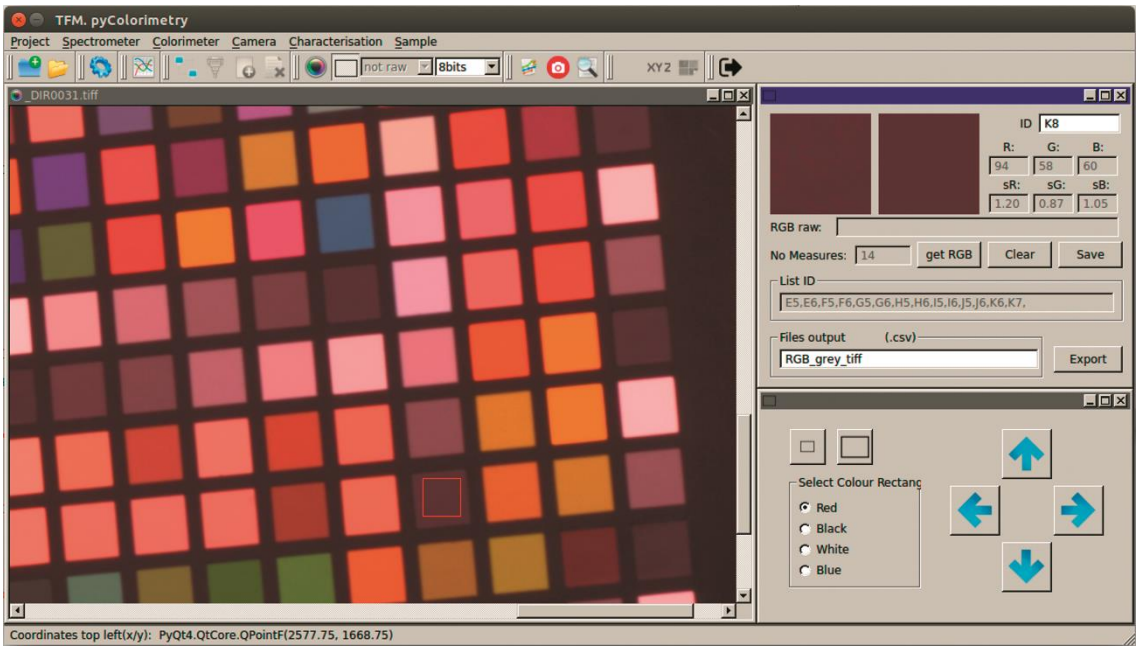


FIGURE 2 Measuring RGB data with pyColourimetry software

3.2 Data processing workflow

For the characterisation of the digital camera, we transformed the original data in the device dependent RGB space into the physically based CIE XYZ colour space. Once the transformation equations are determined, we can represent the image in the sRGB colour space.

In the methodological process, we considered the basic colourimetric recommendations published in the CIE report [16]. Some fundamental aspects described in the report are the usage of illuminants, observers, reference standards for reflectance, viewing conditions, lighting, calculation of tristimulus values, chromaticity coordinates, colour spaces, colour differences and auxiliary formulas.

Once the initial datasets are prepared, we differentiate two stages in the procedure. The first stage aims specifically at the characterisation of the camera, which gives the coefficients of the polynomial transformation equations between the RGB and CIE XYZ spaces. In the second stage, the transformation equations are applied on the input image to obtain the definitive sRGB image (Figure 3).

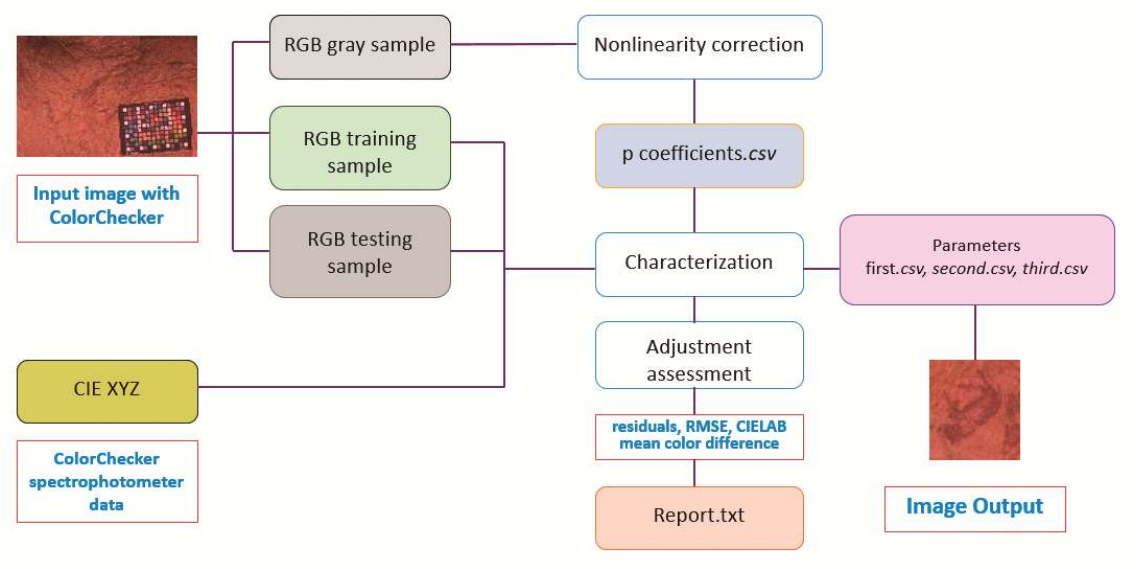


FIGURE 3. Flow chart to show the characterisation implemented for digital cameras.

In the process, we need three sets of samples with RGB data from the selected colour patches: (1) a grey sample to obtain the coefficients for the nonlinearity correction; (2) a training sample to determine the transformation equations; and (3) a testing sample, which will allow us to evaluate the quality of the applied adjustments [12].

The first dataset contains the 15 grey scale patches. The p coefficients for the non-linearity correction are calculated from this first sample [12]. After checking for linearity, the training sample data set is used to determine the characterisation equations. A total of 54 patches of the most characteristic colours of the scene are selected to set the system of equations. The training dataset provides sufficient redundancies for the least squares adjustment.

Finally, the equations are applied on the third testing sample dataset. The software provides statistical estimators and residuals to carry out the quality assessment of the applied adjustments and select the optimum coefficients for the characterisation. This information is essential to make right decisions during the camera characterisation process.

3.3 First stage: Characterisation of the digital camera

The characterisation is carried out in two steps. First, the need for nonlinearity correction of the pixel values is analysed. Although the initial response of the CCD sensor is almost linearly related to the intensity of the incident light, it is unlikely that the RGB output of the camera is linearly related to the CIE XYZ tristimulus values of the scene surface. The raw RGB data are transformed by means of a complex sequence of operations to obtain RGB output values, such as preprocessing, linearisation, white balance adjustment, demosaicing and colour transformation [27].

It is recommended to consider a correction for the nonlinearity of the pixels response as a pre-processing stage in the characterisation of cameras according to Eq. 4. The exponential p coefficients are calculated from the relationship between the response of each RGB channel and the luminance (CIE Y trisimulus value), so that the relationship

between the C_i and C_i' values is linear [12]. The formula is:

$$C_i = C_i'^{p} \qquad (4)$$

where C_i' is the camera's response on channel i (red, green and blue); p is the exponential coefficient; and C_i is the pixel value after the application of the correction for channel i. It is worth noting here that raw values present highly linear behaviour and this correction is not required.

In the second step, the coefficients of the RGB to CIE XYZ transformation equations are obtained from the selected patches of the colour chart. The software implements three polynomial models: linear, second and third order [10,12]. The number of coefficients to be determined is 3, 10 and 19, respectively, depending on the degree of the polynomial (Table I).

Table I. – Transformation coefficients matrix (first, second and third order)

1st order	R	G	B																
2nd Order	R	G	B	R^2	G^2	B^2	RG	RB	GB	1									
3rd Order	R	G	B	R^2	G^2	B^2	RG	RB	GB	R^3	G^3	B^3	R^2G	R^2B	G^2R	G^2B	B^2R	B^2G	1

3.4 Second stage: transformation to sRGB space

The coefficients from the previous stage allows the transformation of RGB pixel values into CIE XYZ tristimulus values. However, in order to adhere to current digital colour standards, a final transformation from CIE XYZ to sRGB space is still necessary. This transformation is calculated following technical recommendations from the International Electrotechnical Commission [28].

The matrix formulas to transform into the sRGB space are:

$$\begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix} = \begin{bmatrix} 3{,}2406 & -1{,}5372 & -0{,}4986 \\ -0{,}9689 & 1{,}8758 & 0{,}0415 \\ 0{,}0577 & -0{,}2040 & 1{,}0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad (5)$$

where XYZ are the CIE tristimulus values.

The CIE XYZ values are first transformed to non-linear sR'G'B' values as follows:

If $R_{sRGB}, G_{sRGB}, B_{sRGB} \leq 0{,}0031308$ then

$$R'_{sRGB} = 12{,}92 \cdot R_{sRGB}$$
$$G'_{sRGB} = 12{,}92 \cdot G_{sRGB}$$
$$B'_{sRGB} = 12{,}92 \cdot B_{sRGB}$$

otherwise, that is, if $R_{sRGB}, G_{sRGB}, B_{sRGB} > 0{,}0031308$ then

$$R'_{sRGB} = 1{,}055 \cdot R_{sRGB}^{(1{,}0/2{,}4)} - 0{,}055$$

$$G'_{sRGB} = 1{,}055 \cdot G_{sRGB}{}^{(1{,}0/2{,}4)} - 0{,}055$$
$$B'_{sRGB} = 1{,}055 \cdot B_{sRGB}{}^{(1{,}0/2{,}4)} - 0{,}055$$

Then, the non-linear sR'G'B' values are converted to digital code values. This conversion scales the above sR'G'B' values by using the following equations: (for a black count of 0 and white digital count of 255 for 8-bits encoding).

$$R_{8bit} = 255 \cdot R'_{sRGB}$$
$$G_{8bit} = 255 \cdot G'_{sRGB}$$
$$B_{8bit} = 255 \cdot B'_{sRGB}$$

The outcome of the process is a digital image represented in a device independent, physically based colour space, which can be rendered at maximum quality in devices compatible with the sRGB colour space.

3.5 Software development

Each computation stage described above is integrated in our software pyColourimetry. The computer software consists of a set of modules that allows users to process colourimetric data, obtain the transformation equations for the characterisation of the camera, and apply the equations on the image to yield images in the final sRGB space (Figure 4).

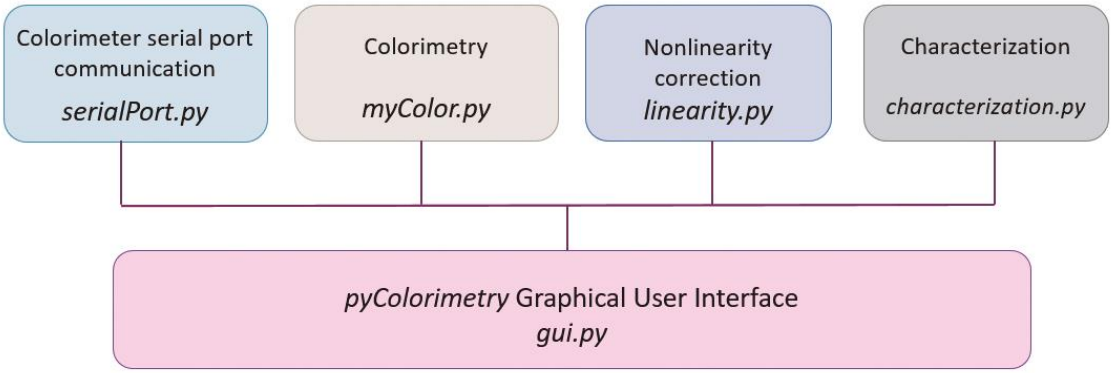


FIGURE 4. pyColourimetry modules

pyColourimetry was developed in Python, which has become a mature language and has all the features required for scientific computing [29,30]. Python is an interpreted, multiplatform programming language, released under a GNU general public license, which guarantees end users the freedom to use, study, share and modify the software. The computer platform chosen for programming tasks was Linux, specifically the Ubuntu 14.04 Desktop distribution [31]. The main functionalities of each module are listed in Table II, and the graphical user interface (GUI) can be seen in Figure 5.

Table II. – Functionalities of the pyColourimetry software

Module	Functionality
serialPort.py	Serial port scanning and detection Communication with the CS-100A colourimeter
myColour.py	Transformation among colour spaces Colour difference transformation Raw data processing
linearity.py	Computation of p linearity exponents
characterisation.py	Setting the characterization parameter
gui.py	Graphical user interface (GUI) Integration of modules User input Generation of reports Generation of sRGB images

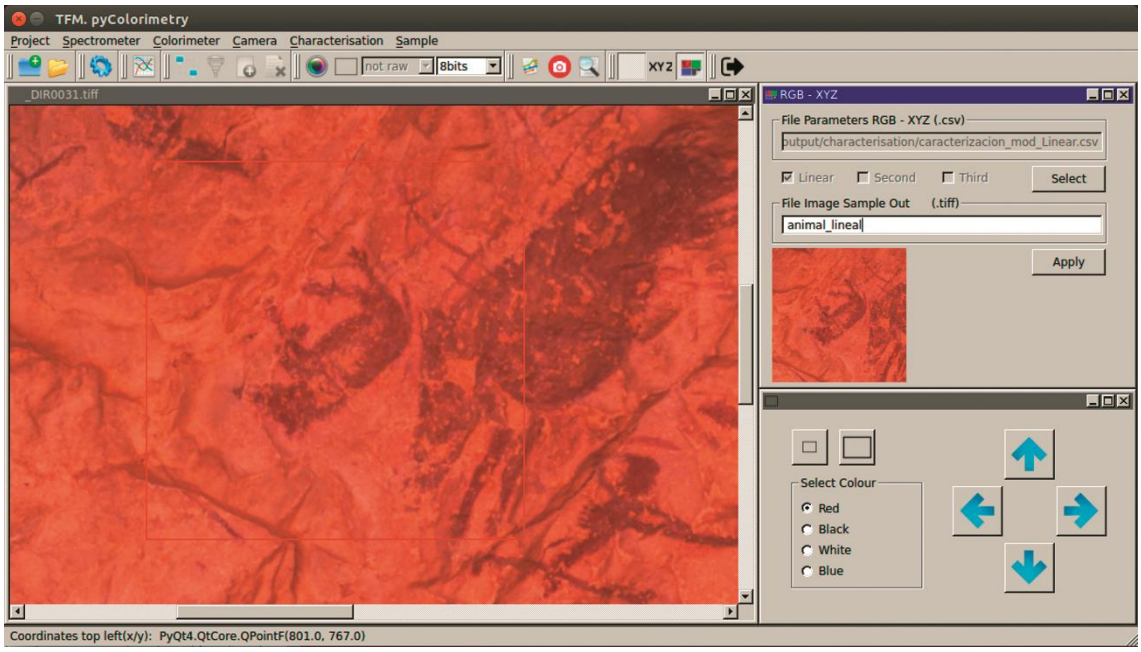


FIGURE 5 pyColourimetry GUI showing a rock art specimen

3.6 RAW data processing

An operating characteristic of modern digital cameras is that the captured raw RGB data are transformed by means of a complex sequence of operations applied by the camera software, such as preprocessing, white balance adjustment, demosaicing and colour transformation [27]. All these operations alter the numerical values of the RGB pixel data used during the characterisation, therefore affecting the calculation of the RGB - CIE XYZ transformation parameters. In our camera, the output of these operations is a tagged image file format (TIFF) picture, which minimises the effects of the internal preprocessing.

A completely different approach for the camera characterisation is to use the raw RGB data instead of the processed TIFF data. This so called raw-based characterisation approach eliminates the influence of those automatic operations from the characterisation process, and provides the advantage of computing the parameters of polynomial models

straight from the sensor response (Figure 6). Although the use of raw data is better for precise characterisation [12], it is uncommon to find software packages with this characteristic. A number of technical complications may explain the lack of such systems, for instance the handling of non-standard file formats or the computational load due to the raw data high dynamic range. In spite of those complications, we decided to write a software package with raw data processing characteristics that eventually gave very good results as explained below.

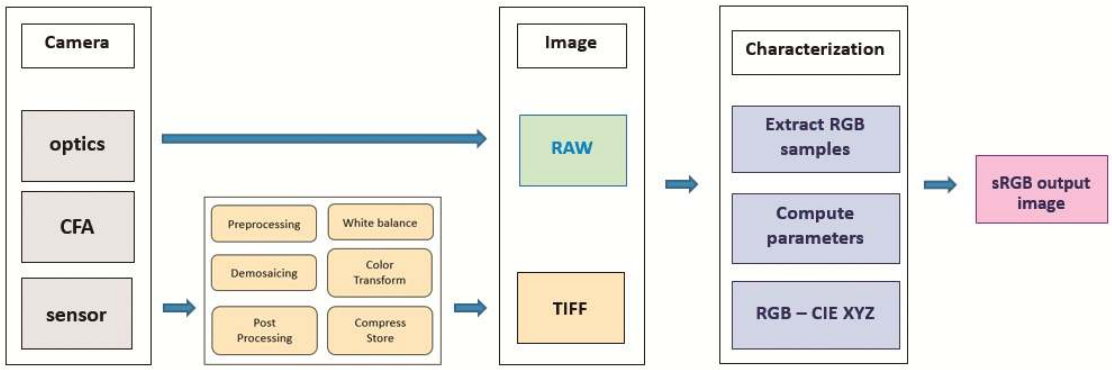


FIGURE 6. Comparative characterisation workflow for RAW and TIFF images.

4. RESULTS

From the sets of training and testing samples, three least squares adjustments using first, second and third order polynomials were conducted (Figure 7). In order to determine the optimal fit for the RGB to CIE XYZ transformation, we analysed both the residuals and mean colour differences. We also provide results for two different processing approaches based on TIFF images and RAW images.

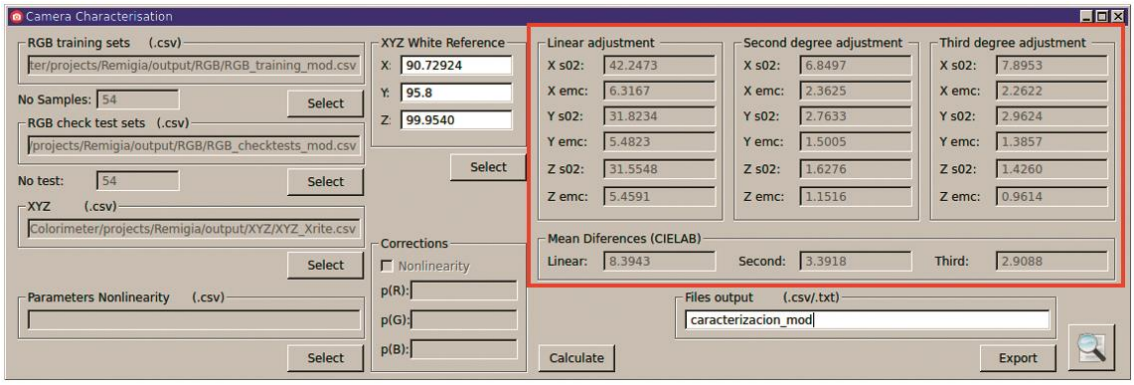


FIGURE 7 Adjustment results for the characterisation of the digital camera (pyColourimetry)

4.1 TIFF data results

Regarding tristimulus values residuals, there are not great differences between the transformations of second and third order. In fact, these transformations gave the smaller residuals (Figure 8). As the degree of the polynomial increases, the statistical and residual results of the adjustment improve markedly. On the other hand, the adjustment with the greatest residuals is that of the linear transformation. Therefore, according to this initial

criterion, the linear transformation should be discarded.

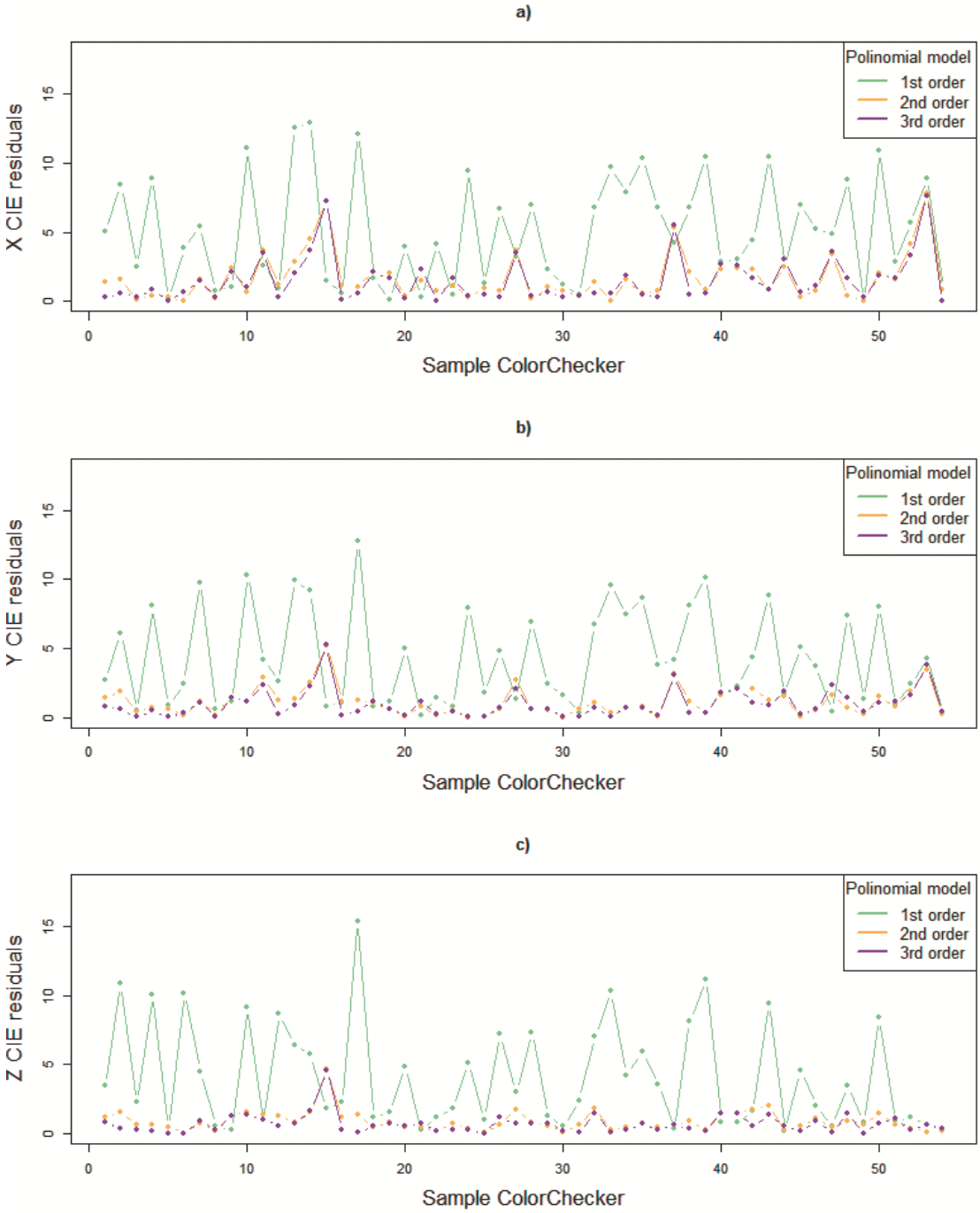


FIGURE 8. Residuals in the CIE XYZ space with TIFF data (1st order; 2nd order; 3rd order): a) X; b) Y; c) Z

The examination of the CIELAB colour differences shows again the contrast between the linear adjustment on the one hand, and the second and third order adjustments on the other (Figure 9). In the linear transform, only a few colour patches are below the acceptable value for ΔE^*_{ab} set at 3 CIELAB units. In accordance with this criterion, the linear polynomial fit is discarded. The second and third order settings yield close values for most patches. It is therefore recommended, as other authors state in previous research, to use either the second or the third order adjustments for the camera characterisation [10,12].

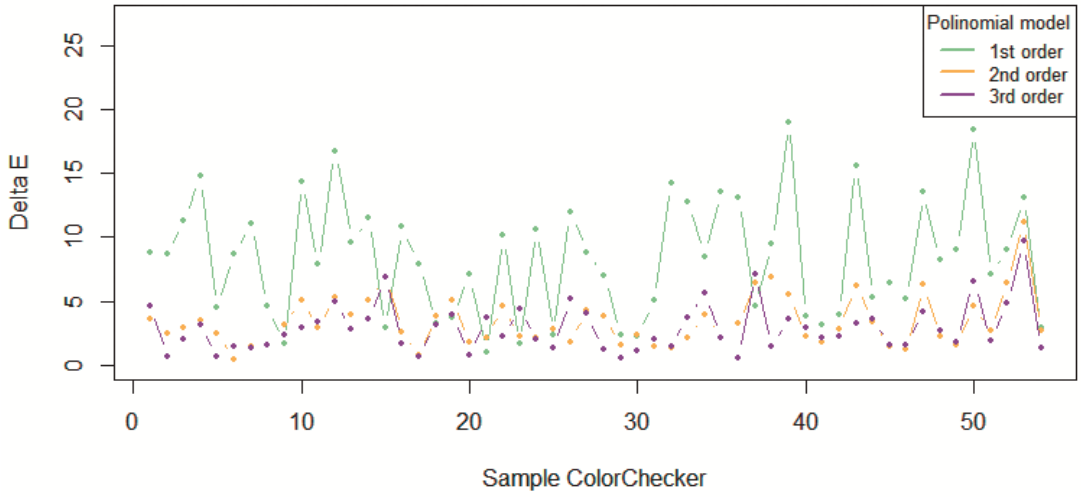
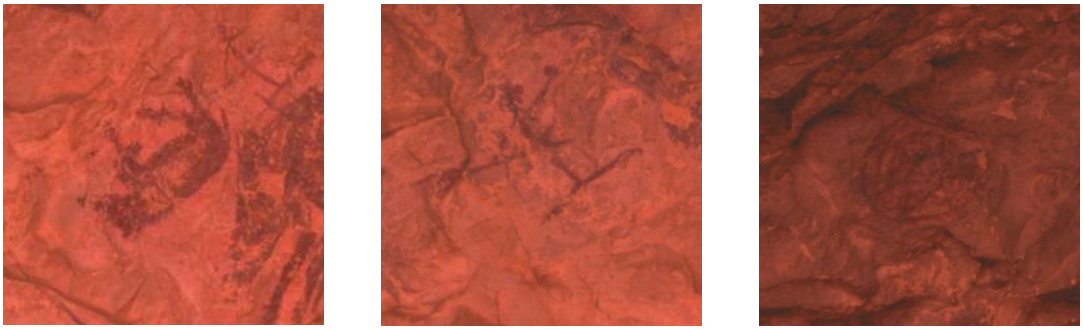


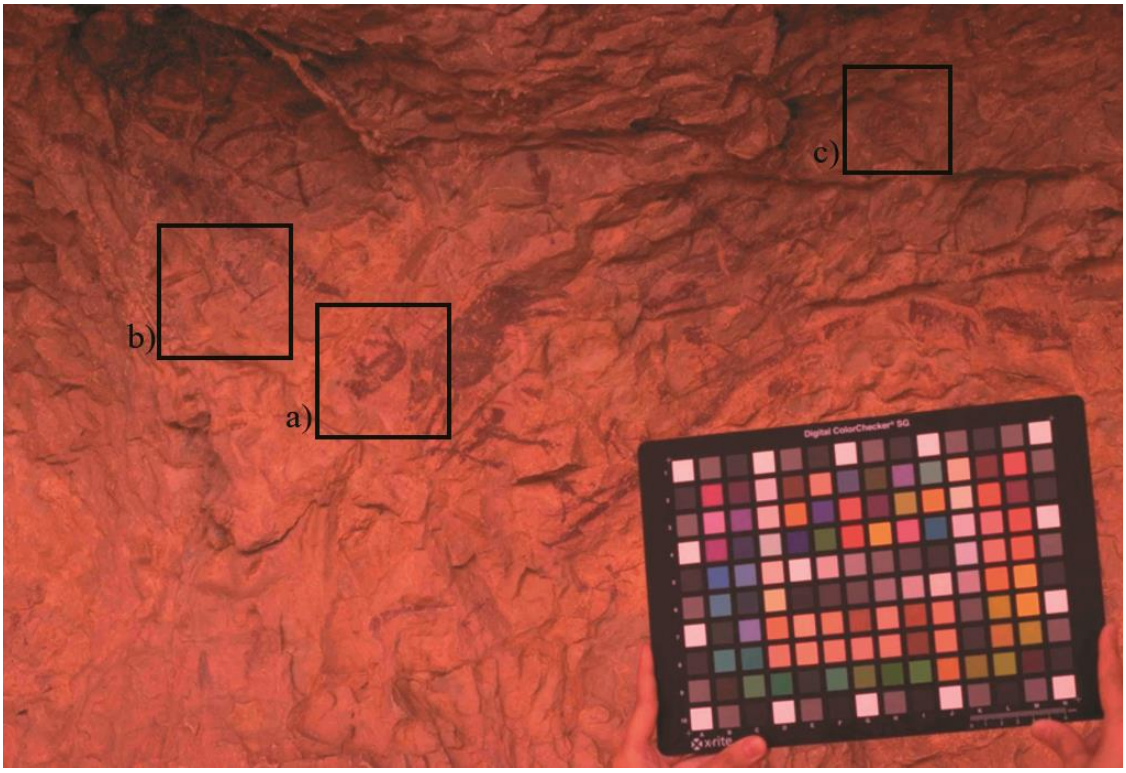
FIGURE 9. ΔE CIELAB Colour differences with TIFF data

As for the colourimetric characterisation of the input image, the transformation equations computed with the three adjustments (first, second and third order) were applied to three specimens that appear in the scene representing a wild boar, a hunter, and a nest [32] (Figure 10). As a result, we obtained output images in the sRGB space for the three transformations (Figures 11 and 14, upper row).

Visual differences among the first order, the second order and the third order transformations are clearly perceived in the first two specimens, the wild boar and the hunter (Figure 11). The best results are obtained with higher order transformations, where the pigment is clearly distinguished from the support. The application of the linear transformation on the three specimen images gives the worst results.

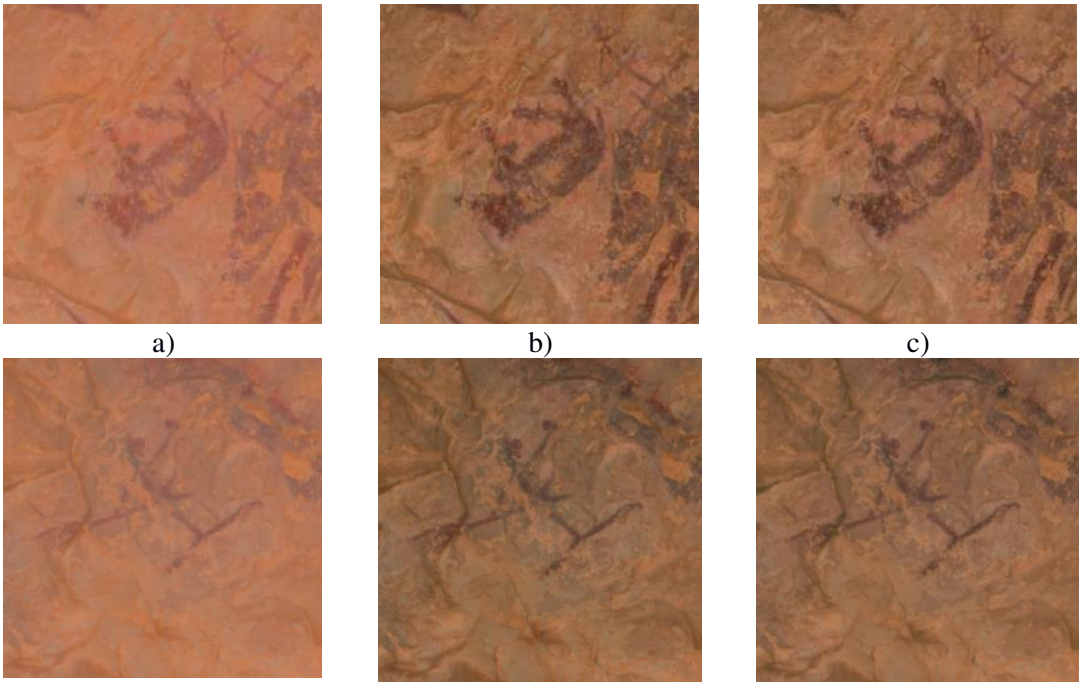


a) b) c)



d)

FIGURE 10 Specimens selected in the scene: a) wild board; b) hunter; c) nest; and d) layout of specimens



a) b) c)

d) e) f)

FIGURE 11 sRGB output images: a, b, c) Animal detail; d, e, f) Hunter detail; a, d) 1st order; b, e) 2nd order; c, f) 3rd order.

4.2 RAW data results

In this section, we give further numeric results and pictures obtained from the adjustments using TIFF data versus raw RGB data (Table III and Figure 12). As in the previous section, the linear model adjusted with TIFF pictures gives large residuals, and therefore this model must be discarded for colour processing. However, the linear model computed from the raw data provides RMSE errors that are half the magnitude of those from the TIFF data.

The second and third order models have different behaviour. In the TIFF data adjustments, the third order model has the minimum residuals, whereas in the raw adjustments, the second order model yields the best result. A comparison between TIFF and raw adjustments shows slightly better results in the TIFF adjustments. According to these figures, a user should apply either the third order model with TIFF pictures or the second order model with RAW pictures, with apparent preference for the former.

Colour differences (ΔE^*_{ab}) follow a similar trend as that observed in XYZ residuals. The linear model gives the highest values, although in the raw adjustment the difference is 4.4 CIELAB units, which is very close to the JND threshold. The colour differences in the second and third order models are reasonably acceptable. Again, a casual user might be tempted to use the third order model with TIFF pictures since it has the lower value (2.9 units). However, this point deserves further discussion (see Section 5 and Figure 14).

Results from raw data adjustments show that the polynomial of higher order does not improve the result (Table III). Instead, the second order polynomial gives the lower colour difference with a similar value to that of the third order polynomial with TIFF. The advantage of using the second order model is the reduction of the order of the polynomial. In this line, authors recommend using polynomials of first or second order, against polynomials of higher order [12,14].

Table III. Comparative adjustment results for the characterisation of the digital camera with TIFF and RAW data

Standard deviation	TIFF picture			Standard deviation	RAW picture		
	Linear	Second	Third		Linear	Second	Third
X	6.3167	2.3625	2.2622	X	3.1098	1.8692	2.2882
Y	5.4823	1.5005	1.3857	Y	2.5051	1.4779	1.8963
Z	5.4591	1.1516	0.9614	Z	2.4153	1.3640	1.9415
ΔE^*_{ab}	8.3943	3.3918	2.9088	ΔE^*_{ab}	4.4079	3.0201	3.6570

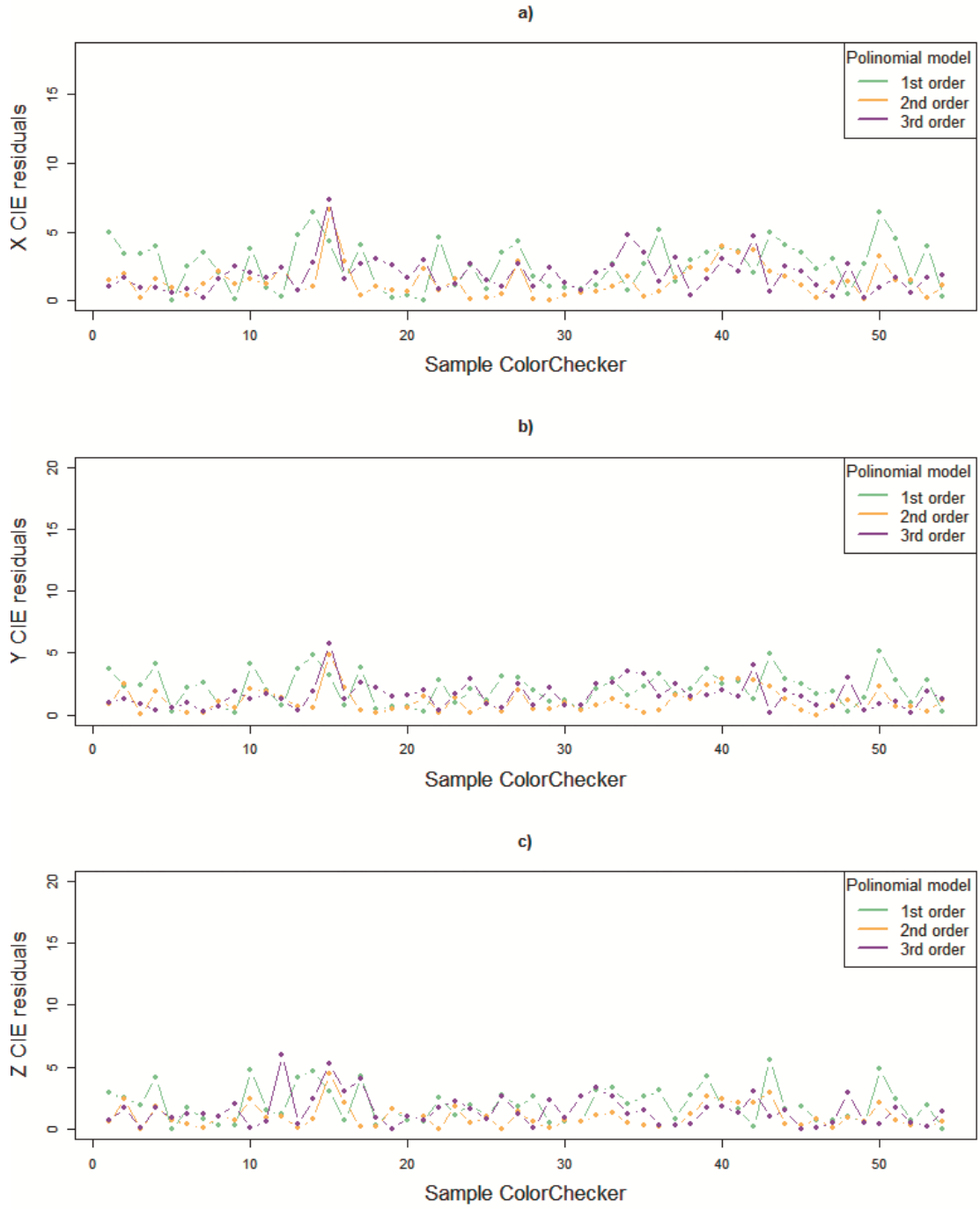


FIGURE 12 Residuals in the CIE XYZ space with RAW data (1st order; 2nd order; 3rd order): a) X; b) Y; c) Z

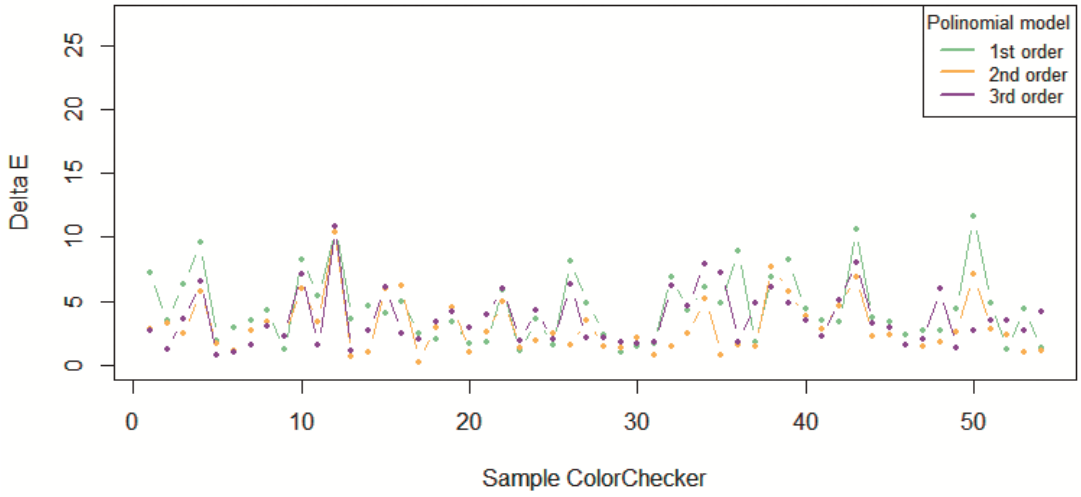


FIGURE 13. ΔE CIELAB Colour differences with RAW data

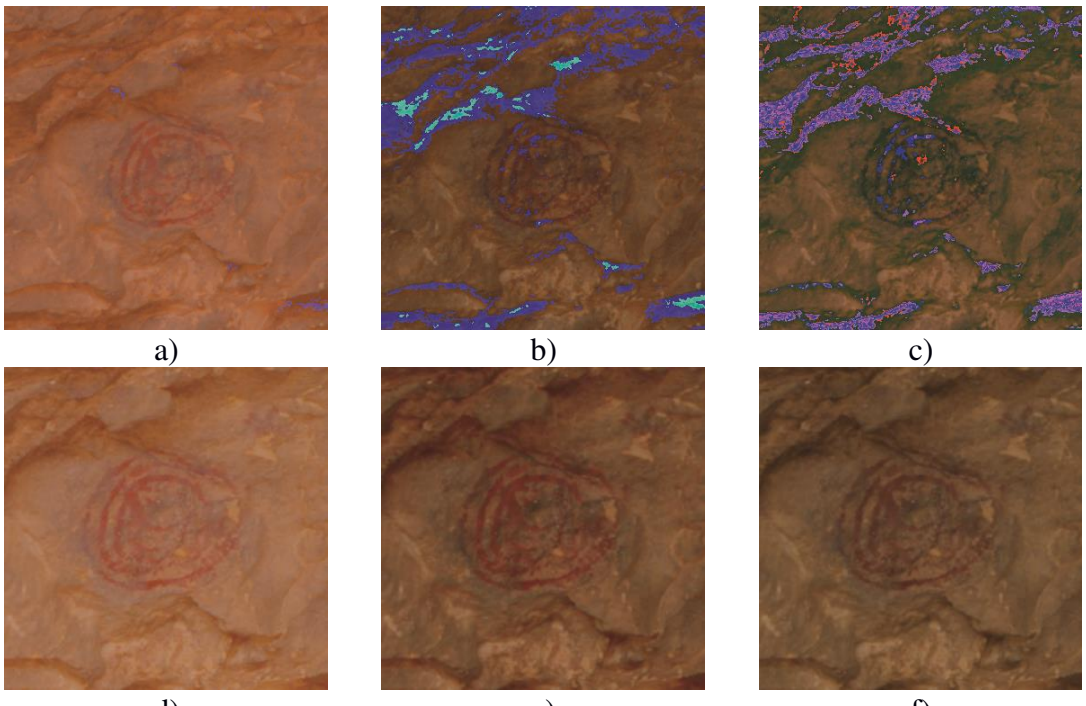


FIGURE 14. sRGB output images of the nest detail: a, b, c) TIFF data; d, e, f) RAW data ; a, d) 1st order; b, e) 2nd order; c, f) 3rd order.

5. DISCUSSION

The processing of TIFF files gives mean colour differences ΔE^*_{ab} in CIELAB units of 8.4 for the linear model, 3.4 for the second order model, and 2.9 for the third order model. It is clear that colour differences decrease as we increase the degree of the polynomial. The worst result is obtained in the linear adjustment. In this case the value of 3 CIELAB units established as tolerance in terms of 'just noticeable difference' (JND) is exceeded, so the linear model should be discarded since it does not adequately describe the relationship

between the device RGB colour space and CIE XYZ colour space. The best results are obtained with the third order polynomial, where the mean colour difference is less than three CIELAB units.

It is quite interesting to consider the results obtained in the third specimen, the nest. While the linear transformation might be thought to be discarded, after checking the output images obtained for both the wild boar and the hunter, in the latter specimen, it seems clear that the best results are obtained with the linear transformation. Despite the high residual value and mean ΔE^*_{ab} colour differences, the final sRGB image obtained for the nest with the linear model is better (Figure 14.a). A smaller number of saturated pixels appears, and the colour obtained after the characterisation is closer to the colour observed in the field site. In the images obtained with either the second or the third order polynomial transformations the number of saturated pixels is substantial. With this particular example, we proved that an increase in the order of the polynomial gave lower adjustment errors, but also created saturated sRGB output images that became useless for colour documentation purposes.

It should be noted that the nest is in the upper-right area of the scene, which is less illuminated due to the prominent relief and curvature at the roof of the shelter. On the contrary, the central area of the image containing the other specimens is nearly vertical and flat. Obviously, the geometry of the cave affects lighting conditions, and this modifies fundamental colourimetric factors such as the reflectivity and colour appearance of the object. In addition, the colour chart was placed on the bottom-right area of the image that corresponds with higher illumination conditions.

The coefficients determined for the transformation equations with the TIFF data were therefore adequate for imaging areas with lighting conditions similar to those of the area where the colour chart was placed, that is, the area close to the wild boar and hunter specimens. In dimly lit areas, as it is the case of the upper side of the image containing the nest specimen, these equations do not function properly.

The most interesting finding of this study was the high performance of the camera characterisation when using raw data. In general, the models created from raw data gave the lower residuals in terms of both XYZ residuals and colour differences. However, the true improvement of raw-based characterisation lays in the quality of the output images. The sRGB images created with the raw-based characterisation contain colours that remind one the true colours at the archaeological site. Moreover, the process proved to be very robust, and worked well with specimens imaged under illumination conditions different to those of the colour chart.

6. CONCLUSION

The results presented in this paper confirm that the initial objective of establishing a methodological process for the characterisation of digital cameras in archaeological research was achieved. It requires the joint processing of both radiometric and colourimetric data. The results are satisfactory and very promising for proper colour documentation in rock art studies.

The proposed workflow for the characterisation of digital cameras is adequate and takes into account the most important technical colourimetric aspects such as illuminants,

standard observers, calculation of tristimulus values, chromaticity coordinates, colour spaces, colour differences and auxiliary formulas.

According the results of the study, we recommend applying the raw-based characterisation with second order polynomial equations to transform data from the camera raw RGB space to the CIE XYZ colour space. The usage of a linear or third order transformation should be restricted to very specific cases due to the higher residual values achieved in the adjustments.

One of such cases may be the processing of image features with illumination conditions different to those in the colour chart area. Although these cases are rare and must be avoided in common practice, researchers must deal with them occasionally. Careful use of artificial light sources and detailed photographic shots containing the colour chart next to the rock art specimen are good practices to avoid troublesome scenarios.

The computer application developed in this study, pyColourimetry, covers all steps of the methodological process, from the measurement and processing of colourimetric samples, up to the characterisation and creation of the output image in the sRGB space. Moreover, our software gives the user full control on the overall computation process.

Future research includes the processing of raw data from additional sensors and cameras with specific file formats, the optimisation of the code implemented for the development of the pyColourimetry toolbox, as well as the inclusion of robust statistical models to obtain improved transformation equations for the characterisation of digital cameras.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support from the Spanish *Ministerio de Economía y Competitividad* to the project HAR2014-59873-R. The authors would like also to acknowledge the comments from the colleagues at the Photogrammetry & Laser Scanning Research Group (GIFLE) and the fruitful discussions provided by Archaeologist Dr. Esther López-Montalvo.

REFERENCES

1. Ruiz JF, Pereira J. The colours of rock art. Analysis of colour recording and communication systems in rock art research. J Archaeol Sci. 2014;50(1):338-349. doi:10.1016/j.jas.2014.06.023.
2. Brill MH. Do Tristimulus Values have Units? Color Res Appl. 1996;21(4):310-313. doi:10.1002/col.5080210404.
3. Lang H. How much physics does colorimetry really need? Color Res Appl. 1997;22(3):212-215. doi:10.1002/(SICI)1520-6378(199706)22:3<212::AID-COL11>3.0.CO;2-G.
4. Higuchi R, Suzuki T, Shibata M, Taniguchi Y, Gülyaz M. Digital non-metric image-based documentation for the preservation and restoration of mural paintings: the case of the Üzümlü Rock-hewn Church, Turkey. Virtual Archaeol Rev. 2016;7(14):31. doi:10.4995/var.2016.4241.
5. Boyd CE, Marín FM, Goodmaster C, Johnson A, Castaneda A, Dwyer B. Digital Documentation and the Archaeology of the Lower Pecos Canyonlands. Virtual Archaeol Rev. 2012;3(5):98-103.

6. Bergman TJ, Beehner JC. A simple method for measuring colour in wild animals: Validation and use on chest patch colour in geladas (Theropithecus gelada). Biol J Linn Soc. 2008;94(2):231-240. doi:10.1111/j.1095-8312.2008.00981.x.
7. Malacara D. Color vision and colorimetry: Theory and applications. Color Res Appl. 2003;28(1):77-78. doi:10.1002/col.10118.
8. McDonald R, Rigg B, Colourists. S of D and. Colour Physics for Industry. Society of Dyers and Colourists; 1997.
9. Ohta N, Robertson AR. Colorimetry : Fundamentals and Applications. England: J. Wiley; 2005.
10. Balasubramanian R. Device Characterization, In Digital Color Imaging Handbook, Chapter 5. CRC Press; 2003.
11. Martínez-Verdú F, Pujol J, Capilla P. Characterization of a Digital Camera as an Absolute Tristimulus. J imaging Sci Technol. 2003;47(4):279-295. doi:10.1117/12.474876.
12. Westland S, Ripamonti C, Cheung V. Computational Colour Science Using MATLAB®. Chichester, UK: John Wiley & Sons, Ltd; 2012. doi:10.1002/9780470710890.
13. Vrhel MJ, Trussell HJ. Color correction using principal components. Color Res Appl. 1992;17(5):328-338. doi:10.1002/col.5080170507.
14. Cheung V, Westland S, Connah D, Ripamonti C. A comparative study of the characterisation of colour cameras by means of neural networks and polynomial transforms. Color Technol. 2004;120(1):19-25. doi:10.1111/j.1478-4408.2004.tb00201.x.
15. Domingo I, Villaverde V, López-Montalvo E, Lerma JL, Cabrelles M. Latest developments in rock art recording: towards an integral documentation of Levantine rock art sites combining 2D and 3D recording techniques. J Archaeol Sci. 2013;40(4):1879-1889. doi:10.1016/j.jas.2012.11.024.
16. CIE. Colorimetry. Commission Internationale de l'Eclairage; 2004.
17. Luo MR. Applying colour science in colour design. Opt Laser Technol. 2006;38(4):392-398. doi:10.1016/j.optlastec.2005.06.025.
18. International Organization for Standardization. ISO 12647-2. Graphic technology — Process control for the production of half-tone colour separations, proof and production prints —. 2004.
19. Mahy M, Van Eycken L, Oosterlinck A. Evaluation of Uniform Color Spaces Developed after the Adoption of CIELAB and CIELUV. Color Res Appl. 1994;19(2):105-121. doi:10.1111/j.1520-6378.1994.tb00070.x.
20. FADGI. Federal Agencies Digital Guidelines Initiative. http://www.digitizationguidelines.gov/.
21. Files M, Griggs K. Technical Guidelines for Digitizing Cultural Heritage Materials : Creation of Raster Image. Image Rochester NY. 2016.
22. Bureau Metamorfoze. metamorfoze. https://www.metamorfoze.nl/english.
23. van Dormolen H. Metamorfoze Preservation Imaging Guidelines. 2012;(2).
24. Korytkowski P, Olejnik-Krugly A. Precise capture of colors in cultural heritage digitization. Color Res Appl. 2017;42(3):333-336. doi:10.1002/col.22092.
25. CIE. Colorimetry – Part 6: CIEDE2000 Colour-Difference Formula. 2013.
26. UNESCO World Heritage Centre. Rock Art of the Mediterranean Basin on the Iberian Peninsula. http://whc.unesco.org/en/list/874.
27. Ramanath R, Snyder WE, Yoo Y, Drew MS. Color image processing pipeline. Signal Process Mag IEEE. 2005;22(1):34-43. doi:10.1109/MSP.2005.1407713.
28. International Electrotechnical Commission. IEC/4WD 61966-2-1: Colour Measurement and Management in Multimedia Systems and Equipment - Part 2-1: Default RGB Colour Space - sRGB. Int Electrotech Comm Tech Rep. 1998.
29. Oliphant TE. Python for Scientific Computing. Comput Sci Eng. 2007;9(3):10-20. doi:10.1109/MCSE.2007.58.
30. Rossum G Van, Drake FL. The Python Library Reference. October. 2010:1-1144.
31. Ubuntu. Ubuntu operating system. https://www.ubuntu.com/.
32. Sarriá Boscovich E. The rock printigns of Cova Remigia (Ares del Maeste, Castellón). Lucentum An la Univ Alicant Prehist Arqueol e Hist Antig. 1988;(7):7-34.