Received June 30, 2020, accepted August 6, 2020, date of publication August 17, 2020, date of current version October 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3017390

An Adaptive Colour Calibration for Crowdsourced Images in Heritage Preservation Science

MIGUEL ANTONIO BARBERO-ÁLVAREZ, JOSÉ MANUEL MENÉNDEZ, (Senior Member, IEEE), AND JUAN ANTONIO RODRIGO

GATV-SSR, Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Corresponding author: Miguel Antonio Barbero-Álvarez (mba@gatv.ssr.upm.es)

This work was supported in part by the project MIPAC-CM (*Monitorización por procesado de imagen y ciencia ciudadana para la conservación de materiales del patrimonio cultural*—Image processing monitoring and city science for maintenance of cultural heritage materials), under Project Code 2018/NMT-4913, in colaboration with the Centro Nacional de Investigaciones Metalúrgicas (CENIM), Centro Superior de Investigaciones Científicas (CSIC).

ABSTRACT In this paper, an adaptive image calibration technique suitable for cultural heritage surveillance in controlled environments is proposed. Heritage pieces, as objects important to conserve, are prone to degradation and thus in the need of a constant watch in order to assess the moment where maintenance is required, preferably in a non-invasive way. By means of a crowdsourcing approach, we take as surveillance cues the pictures visitors take of the piece under surveillance to avoid the need of a permanent human effort on monitoring. In this paper we propose a novel image calibration approach destined to monitor the state of a cultural heritage piece taking as information cues photographs of different sources. Our technique aims to surpass the colorimetry difference between pictures taken by different cameras, performing an adaptive linear interpolation having a given display with a set of colours as reference values, thus transforming the initial colour space of any picture to a fixed one. This is performed in the RGB domain. Lastly, the pictures are transformed to the CIELAB space to extract the colour coordinates and measure the overall objective quality of the calibration. Acting like this allows the heritage monitoring of its possible degradation along time in an effective way without incurring in expensive resources.

INDEX TERMS Heritage science, crowdsourcing, cultural heritage preservation, digital image processing, calibration, colour science.

I. INTRODUCTION

Cultural heritage, as a concept, is minted by UNESCO as the cultural attributes that are inherited from the past to our current society [1]. In recent times people have grown conscious of the importance of our cultural heritage and duties of safeguarding responsibility have been acknowledged. Important movements have been formed around this aim, establishing codes and rules on an international level. The progressive digitalisation of society and growing interest in multimedia content has allowed to develop the heritage preservation science [2], [3]. Due to the visual nature of monuments and objects, the multimedia information technology is widely employed. A huge fraction of the world's population is in possession of a cellphone, which enables the existence of a global array of information sources; and is considered today as a potential analytical instrument [4], [5]. Recent papers have been observing the possible benefits of gathering information by crowdsourcing on heritage science [6], [7]. Amongst the different approaches and orientations, the collection of user-generated information and pictures is included [8], [9], which is interesting to ponder enabling on any object susceptible of deterioration and, thus, with a need for constant surveillance and maintenance [10]. Colour is an indicator of the chemical and physical status of matter [11]. Many kinds of matter change colour on their surface when they are under corruption or in damaging environments [12], for instance under oxidation or corrosion cases [13], [14], so visual and colour information can give important cues about when to apply maintenance on heritage pieces.

Cameras are built in correspondence to the human visual system [15], [16] and digital pictures that represent the acquired visual information are usually to be found represented in the RGB colour space [17]. Colour and its representations are an important source of visual information in pictures and video. When visitors pay visit to a heritage piece,

The associate editor coordinating the review of this manuscript and approving it for publication was You Yang.

the pictures they may take on its installation (even without flash) can be used by maintenance employers as cues to get information on its state as part of the Preventive Maintenance Plan without a need to personally control everything by themselves, which can be expensive in time [8]. Nevertheless, due to differences in fabrication or physical characteristics of the components of the acquisition devices, the images resulting by photographing the same scene with independent cameras may look different to the same human [18]. If during a visit every obtainable picture is taken by a different visitor, it is assumed that they will be originated from different cameras, and thus the resulting visual information may be different even if taken at the same moment and under the same lighting conditions. Thus, an homogenisation of the visual information, such as colour, is needed in order to perform the maintenance task correctly [17]. Then employees could trust the information regardless of its origin and effectively recognize when some piece needs to be taken care of by judging on the supplied data. Since here it is assumed that camera calibration cannot be performed due to impossibility to access the devices of third parties, it should be done after the acquisition and before the information storage.

Hereby, this paper aims to expose an improving focus on image processing to adapt crowdsourced information for cultural heritage maintenance in controlled environments. Previously published papers practiced photography processes without any further operation, or a simple white balancing, and they threw a study on the variability and stability of the information between images of different cameras by taking difference and tolerance measures [10] of the CIELAB colour coordinates of the pieces under observation. CIELAB is widely employed as an accurate colour specification space in heritage science [14], [19]. Our work proposes adaptive operations of global calibration that aim to improve the performance by homogenising any set of crowdsourced pictures acquired of a defined display, consisting of a colour checkerboard and indicating material tags, located next to the heritage piece under observation in a controlled lighting environment. In a similar context to [20], the ending color coordinates of the material tags in CIELAB are then extracted as a means of obtaining their chromatic information. We base the evaluation of the quality of our results on the measures of recent works of analogous approach like [10] and [21], and our results show that our approach improves the stability and quality of the information for future real applications. Effectively, our proposal is thus suitable for recognizing colour changes on matter, and thus its degradation, alongside time and judging the necessity of maintenance of the piece in question in a non-invasive way. Related approaches [14] show that with a single valid retrieved colour cue on daily or weekly basis it should be enough for an effective surveillance.

The following text is thus structured in this way: In the second section a state-of-art is outlined. The third section throws an outline and a chain of operations that shed light on the practical application of colorimetry for casual users, seeking easiness of understanding. The information given culminates in our proposal; the fourth section presents the experimental scenario and its procedure, described in a degree of detail and justifying the taken decisions. In the last section, the results are explained and their implications discussed.

II. RELATED WORKS

Under today's scope, Cultural Heritage Preservation has become a hot topic. The existing national or global initiatives see in it an array of opportunities regarding it. The digital horizon plays a big role in this circumstance. The administration of today's countries and political entities regard the promotion of cultural heritage preservation as a part of their activities. There exist initiatives in a national level in many parts of the world [7], [22] that consider technology and the digital conjuncture as an opportunity to ease heritage preservation in different manners and, alongside, draw the interest of the general public to it. Even entities like the European Commission have taken these considerations into account, bringing Cultural Heritage Preservation as a part of their Horizon 2020 program (the EU Framework Program for Research and Innovation) due to its potential in research possibilities [8], [23]. There exist several kinds of proposed state-of-art focuses to obtain relevant data for the prosecuted aims [14]. Crowdsourcing is the most interesting due to its potential for scientific researches [24], [25], spanning different approaches for the possible distinct aspects of cultural heritage preservation [6].

Monitoring technology, e.g. surveillance of an object by means of capturing devices, can be introduced as a part of heritage science by adding new features to improve its process. A traditional monitoring is performed by means of strategically placed devices. However, due to the technology improvements, it is possible to employ data from third parties, such as visitors to the heritage site, to help surveillance [9], [14]. This can be used in combination with the system installed by the surveillance organism [26].

When summoning third parties for using their supplied information for heritage maintenance, its data can be of any nature, from writings to statistics, opinions and even taken images or measured ambient conditions [7], [8]. Involving people in the preservation of cultural heritage is a method with more recurrence every time [19]. By involving people in the task it may serve as a stimulus to pique their interest towards the object in question [27], supported by the easiness to obtain information, and help them feel as an important part of the preservation process [9], [28]. Gamification schemes and dedicated crowdsourcing platforms which enhance user experience and interactivity while collecting important data for preservation and monitoring are likely to be attractive to third parties, and thus helpful for the heritage maintenance task [3], [8]. Frequently, those platforms also include a processing system of variable complexity destined to analyse the gathered data for the desired usages and goals [3], [9].

In the current times there has been some research with successful experiments involving crowdsourced information destined to help preservation [5], especially regarding control

and surveillance [20]. This means that, rather that doing traditional preventive maintenance, having a supply of information coming from third parties helps having a constant knowledge of the status of a heritage piece; this can lead to discarding worries and preventive maintenance resources in favour of knowing the moment of repairing.

Of the supplied crowdsourced information, image data are likely to be helpful when dealing with heritage preservation [29]. Amongst all the kinds of image data available, information related to colour and colorimetry is been regarded as significant when inspecting art pieces and objects exposed in museums, as a non invasive method [30]. Standards for conservation of works of art, in fact, incorporate surface colorimetric analysis to acknowledge potential changes in a chemical composition level of pieces of interest [12], due to the science involved behind it [31], [32]. Research works involving third party colorimetric data for monitoring of exhibition pieces and monuments have been published, supporting the interest in this field [13].

Mainly, the considered research works employ crowdsourced data from cellphones of third-party users as supply for the control task [10], [29]. There, they assess that taking pictures of the visitors to a heritage piece and extracting colorimetric information after performing some needed preprocesses is a satisfactory manner of evaluating the status of a monitored object, while being a non-invasive approach at the same time [20]. They also state the importance of CIELAB colorimetric information for their purposes and the observation of employing simple, minimal calibration processes [19]. Whilst calibration of the images is necessary in order to start from a common ground after obtaining images with different characteristics from different sources, an excessive calibration process may result in a hurdle to achieve an exact, accurate result due to a potential addition to random chemical uncertainty of chromatic change in objects.

In these works the final CIELAB coordinates are bound to be used as objective measures. Since the image information is obtained there by objective physical means, e.g. spectrometers, the analysis of the results must be coherent in a way of being assumed as absolute and unbiased. However, as visual information is perceptual, there exist different approaches for camera homogenisation that evaluate their results employing measures taking perception into account, such as SSIM or the S-CIELAB model (named by some authors *Color Image Difference – CID*), after performing the needed corrections between samples [21], [33]. Depending on the scientific application either one of the approaches is correct. Concretely, for the need of a surveillance alongside a period of time where a given reference is the subject of comparison, an objective, general measure and its evolution are more suitable to be observed. Works with a similar approach assure the importance of objective measures [34].

Effectively and taking this into account, in this paper we propose an alternative to the works described in [10] and [29]. It also acts as a starting point where further research and improvement based on more refined mathematical processes can take place with ease in a near future.

III. DEFINITION OF CONCEPTS

A. COLOUR SPACE IN THEORETICAL TERMS

Pictures, as sources of colorimetric information that have to correspond with the human visual perception, are normally found represented in colour spaces that match it. The CIE1931 RGB colour space [35] is a mathematical convention for representing colour in a manner faithful to the human perception and with a certain ease of operation. It describes, therefore, a three-dimensional Cartesian-like space, in which each possible colour stimulus, defined as a point in said space, could be composed as a weighted sum of the different red, green and blue matching coordinates.

Also Cartesian-like is CIE1931 XYZ. XYZ can be described as an analytical and intermediate colour space but not a representation-oriented one. Its generic nature makes it device-independent and a common ground for calculations and modifications on colour without the bias and limitations technology and perception impose to RGB. It is considered a ''master'' space, since it is directly derived from the equations describing the physical colour spectre by matching stimuli, and other spaces can be deduced out of it. It is also based in three stimuli due to the initial mathematical conception of light perception, but unlike RGB, it's not a perceptual colour space. The X, Y and Z stimuli are matching values, that describe any possible colour, lightness information included. Even existing but not perceivable colour tones (imaginary colours) can be described within this space. Of its three stimuli and components, Y is a photometric quantity and represents lightness information or luminance, while X and Z are imaginary, parametric derivations of red, green and blue whose values depend on Y's [35]. The equations to move between RGB and XYZ spaces are detailed in section III-C.

Due to its variable parametric nature, the shape of the visible range in CIE1931 XYZ space has no clear, recognisable general appearance in its Cartesian representation. Nevertheless, by establishing certain values for the adequate parameters, a solid-like representation can be achieved where the array of visible colours can be found and classified. Those are effectively shared with other related spaces, including RGB. The XYZ coordinates are analytically related between them. Therefore, a simpler representation can be achieved if one of the coordinates, like Z, is described as a composition of the other two as $z = 1 - x - y$ (being all the terms the normalized coordinates with respect to the total sum), effectively reaching a two-dimensional space. This variant is called the *xy space*, where x and y are normalizations of XY given a specific luminance value Y. x and y are also called chromaticity coordinates, and this two-dimensional planar space defines the perceptual colour information for any possible colour tone isolated from luminance. This is commonly known as the CIE xy colour plane or chromaticity plane, and it serves as basis for the further developed colour spaces.

When luminance information is desired, the xyY colour space can be built. In this representation, the plane is expanded as a three-dimensional non-regular solid that locates in space the all possible chromaticity and luminance combinations for the human perception (figure 1 Left).

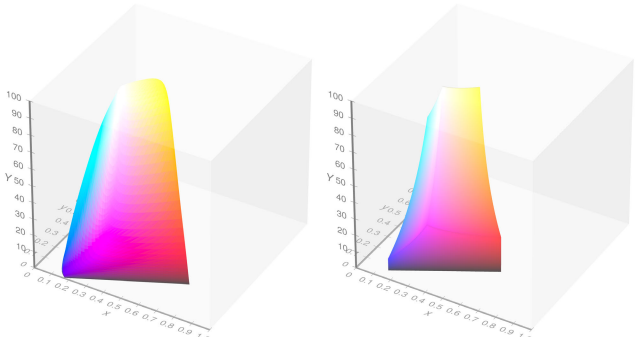


FIGURE 1. Left: xyY solid as an extension of the xy plane. Right: xyY solid of the sRGB gamut, intersection of the complete xyY solid with the sRGB gamut triangle. Source: [36].

The CIE1931 spaces establish the relationship between colours in a non-uniform manner. This means, when the magnitude of the algebraic distance between two colour points in a CIE1931 space is equal to the distance between two other different points in the same space, the human eye may not perceive them as the same. The perceived distance varies depending on the position of the colours in space. Aside, lightness, like the Y-coordinate, does not share a linear relationship with chromaticity, which is also something uniformity should surpass [35], making the desired uniform spaces the most similar ones to an algebraic ideal for managing colour information. The proposed CIE1976 L*a*b* colour space or CIELAB is defined by three orthogonal coordinates calculated out of the XYZ space depending on an established set of reference coordinates for the white point [37]. It was developed with the intention of presenting differences between tones of similar proportion and magnitude throughout the whole domain, so it fulfills the need of uniformity. It has been recommended by CIE as a standard for device-independent space [38]. Within the L*a*b* components, the L value stands for lightness, similar to luminance Y but not representing the same information, ranging between 0 and 100; and the a* and b* colour coordinates, that on paper may fluctuate between −120 and 120, represent orthogonal chromatic ranges between complementary colours [35], [39].

The CIELAB solid of perceivable colour, as described by its orthogonal Cartesian dimensions, can be observed as an irregular polyhedron, whose vertices are the coordinates corresponding to the RGB and YCM primaries and black and white. The vertical axis of this solid is the L scale, and the others connect the complementary primaries (figure 2). Due to the highly non-linear relationship between CIELAB and the perceptual colour spaces, the coordinates for the primaries are concrete real values. Unlike the master space XYZ, CIELAB doesn't conceive imaginary colours because it aims

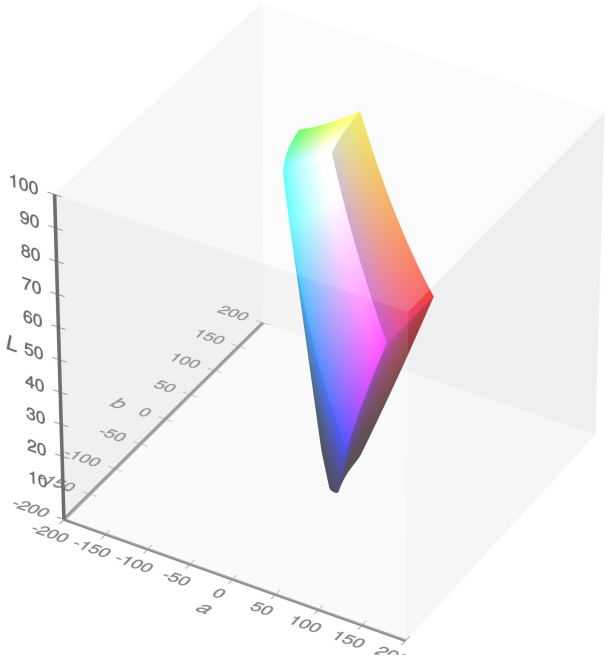


FIGURE 2. Solid of digitally displayable colours in CIELAB space. Source: [36].

to bring uniformity to the perceived ones, so its equations limit every conceivable value outside of the polyhedron to a point on its surface. There also exist colours and RGB combinations humans can perceive that don't exist in CIELAB's domain [40] and thus get clipped to the most similar existing value.

CIELAB's characteristics and independence make it suitable for analytical operations, transmission and processing before converting the data to a more presentation-oriented space. Its usage is especially interesting for calibration tasks, hence its uniformity and independence that help managing data from different sources [17], [41].

B. COLOUR INFORMATION IN TECHNOLOGY

When considering the light (as perceptual RGB values) acquisition of cameras with their optoelectronic block, some restrictions and limitations are to be encountered that hinder the ideal management of the visual information. The main limitation to be encountered is the unavoidable implementation of a gamma operation that is imposed by the standards that manage RGB colour usage by digital devices. Whilst this operation was a necessary step for optimizing the information during the time cathodic ray tube screens were amply used, gamma has continued nowadays as a way to facilitate the acquisition and display of high dynamic ranges images. Its effects are to be fulfilled during the acquisition in the optoelectronic devices. Here, the values the digital device operates with are the real, perceived RGB modified by the imposed gamma. This operation is also found during presentation by displays. The displaying device should be presenting RGB values corresponding to the real, perceivable ones by humans,

so an inverse gamma operation is performed over the digital colour information. The middle step that happens in the device domain, between acquisition and display, forces into a necessary conversion from a physical signal (real-valued) to a discrete digital one. Both effects, the device-induced gamma and the necessary digitisation imply that the explained ideal mathematical theories of colour can't be fully translated to practical cases.

Therefore, the perceived colour information in the physical world is described inside the device with one of the series of proposed RGB digital spaces. The certain chosen discrete space depends usually on one of the digital TV signal standards in use at international level or, equivalently, on the digital TV recommendation followed by the camera technology in usage, since those are the norms that regulate digital colour. The mentioned gamma operation is defined differently for each digital RGB space as an exponential correction for the mathematical RGB values found in real life.

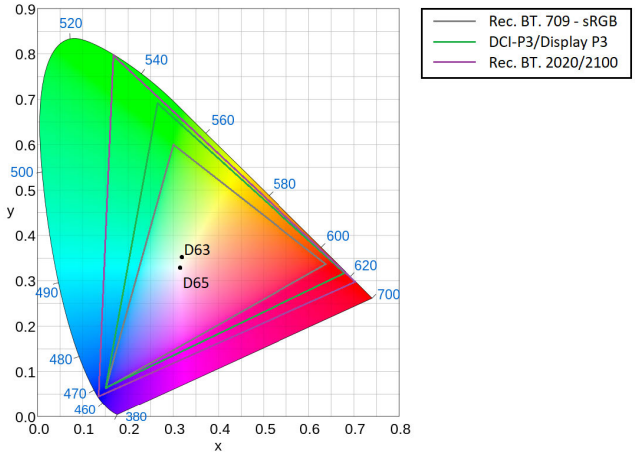


FIGURE 3. CIE chromaticity plane depicting the covered gamuts for sRGB, DCI-P3 and Display P3 and recommendations BT. 2020 and BT. 2100, alongside the white points for the CIE standard illuminants with 2° deviation D65 and D63. Modified from source: [42].

Digital RGB spaces are relative spaces defined by their primaries. Those are specific red, green and blue hues given by defined xy coordinates in the chromaticity plane. A triangle is then formed by connecting those. (figure 3) The colours in the area within form the variety of colour or gamut that can be reached to present in a particular digital RGB space defined by the primaries as vertexes. When representing a gamut in the xyY space, including luminance, it's seen as the intersection of the original xyY solid of the perceivable colours with a prism of triangular base along the Y axis. Due to the chromaticity plane being limited by its particular shape, if it is intended to cover it in its entirety in a single gamut, the xy coordinates for the primaries should in this case correspond to imaginary colours, outside of its range [35] (figure 1 Right).

Any digital RGB space also implies the need of a white point of reference, so it defines the exact white the maximum value of the space takes. Usually for RGB digitisations it is chosen a canonical source of illumination equivalent to D65 with a standard observer of 2° (that is, in front of the light source and with a perpendicular deviation not bigger than 2 degrees), considered as the reference white temperature for digital applications by CIE [18], [43]. Values in (1) represent the xy coordinates for the standard D65 2° observer illuminant. However, other possible white references may be used. The white point lies in the centre of the gamut triangle in the xy plane, as seen in figure 3.

$$x_w = 0.3127, y_w = 0.329 \tag{1}$$

Thus, each image pixel represented in a digital RGB-like colour space is described by a tuple of integer values for red, green and blue; those define its own tone as its spatial position in the gamut. The possible value ranges for each of the three channels, due the necessary discretisation that comes with a digital signal, lies between 0 and 2^n, n being the number of bits/pixel the image is coded with (in each independent channel). Thus, the gamut is discretised in 2^{3n} possible channel combinations.

The different standardized proposals of digital RGB variations differ on the coordinates of their primaries, the area and plane coverage of their gamuts aside of the distinct gamma operations. It can be easily seen that, if any digital RGB space spans in a subset of absolute chromaticity and luminance but they are equally quantified depending on the number of bits defined by the device, digital RGB spaces are thus relative to the device. Their absolute values are then to be found in xyY or XYZ.

The most widespread implemented RGB digitisation scheme is called standard RGB or sRGB, whose primaries are registered in recommendation ITU-R BT.709 [44], adopted as the standard for High Definition - HD signal in TV. This is the commonly selected color space for most of the cell-phones cameras and conventional photographic and video cameras, with some notable exceptions. Its conception relies on the desire to have a standard default colour space every high-end device shares so the communication between them is easy without incurring in expensive forward and backward transformation operation. Its backwards compatibility and its possibility of being shared between devices of different fabrication dates till the most modern ones contribute to that. The gamut defined in ITU-R BT.709 covers 35.9% of the xy plane (figure 3) and it is the most followed recommendation as a consequence of sRGB's prevalence. Accordingly, the gamma operation performed by the optoelectronic signal converter should exactly coincide with the electrical response the CRT monitors had.: $V = L^{0.45}$, with $\gamma = \frac{1}{2.2} = 0.45$, L being the input light signal and V the voltage signal.

However, the practical implementations of sRGB today differ slightly to gather the possibility of handling High Dynamic Range - HDR images according to the human visual system capabilities. In order to increase sensibility in very dark values, its corresponding gamma divides itself in two parts, one for low (where the human eye sensitivity is higher) and other for high values, as it is described by equations (2)

and (3) for computer images [45].

$$\gamma(q) = \frac{q}{12.92}; \forall q = \frac{R}{255}, \frac{G}{255}, \frac{B}{255} \leq 0.04045 \quad (2)$$

$$\gamma(q) = \left(\frac{q + 0.055}{1.055}\right)^{2.4}; \forall q = \frac{R}{255}, \frac{G}{255}, \frac{B}{255} > 0.008856 \quad (3)$$

In order to obtain the ideal RGB values of a digital image in sRGB, it is necessary to reverse the effects of the previously mentioned gamma operation by inverting it.

While Rec. ITU-R BT.709 and sRGB are the most widely employed options for today's generalist technology, the commercialization of screens and displays allowing wider colour gamut imposes new possibilities. Thus, color spaces such as DCI-P3, initially designed for the film industry and its professional screenings, start to be employed [46] (figure 3). Especially of interest is the case for the modern models of cell-phones, fundamentally those from Apple and more recently operated Android ones. Since the cameras and screens, due to the possibility of video watching in any place, are some of the biggest appeals in phones for consumers nowadays, the most modern screens are able to bear an extended color gamut that can match the one employed by the content. Thus, it makes the ability of displaying information of DCI-P3 colour possible in a consumer screen.

DCI-P3 shares with sRGB the coordinates for the blue primary, effectively extending the triangle with the red and green primaries, obtaining a gamut 25% bigger than sRGB's [47]. Its exponential gamma has a value of 2,6 and its white reference resembles more the D63 than the canonical D65 illuminant [48]. However, a variation exists, called Display P3, which uses the sRGB gamma and the D65 white point reference, effectively making it basically a wider gamut sRGB. Even a wider gamut is allowed by the UIT-R BT. 2020 recommendation, developed for Ultra High Definition - UHD TV signal. It covers almost three quarters of the xy plane (75,8%), without including any imaginary colour in it [49]. Another recommendation, ITU-R BT.2100 [50], supports HDR UHD TV signal while utilising the same gamut as ITU-R BT.2020. The featured table in (1) sums up the most basic information for the gamuts of the mentioned spaces to be found in the chromaticity plane.

In this work, sRGB is assumed to be employed due to its prevalence since it suits for basic and moderately advanced devices. While Recommendation ITU-R BT.709 imposes a clipping of the dynamic range, for this application we work with the full range for 8 or 10 bits, from 0 to 255 (in case of 8 bits) or from 0 to 1024 (for 10 bits).

C. TRANSFORMATION BETWEEN COLOR SPACES

The work described in this paper aims to recover exact CIELAB coordinates based on a given set of concrete colours defined by digitalised RGB information acquired by cameras.

The first step of this followed transformation process consists in achieving the absolute values in XYZ. Theoretically, RGB can be transformed to XYZ and vice versa by means of a linear, biyective application. However, the digital-only nature of the available sources of visual information means that this process is expanded: first, it is necessary to convert the available data from spaces such as sRGB to the canonical RGB space by means of its inverse gamma operation, as it is explained in equation (4), and only then perform the linear application that transforms the three red, green and blue coordinates to X, Y and Z.

$$RGB_{lin} = \gamma^{-1}(sRGB) \quad (4)$$

being γ the function described by equations 2 and 3.

Thus, this transformation is made up by two steps: a non-linear and a linear one. This implies that biyectivity, in the very worst of cases, isn't always fully guaranteed due to the unavoidable discretization of the values if it's too broad. RGB is defined by its set of integer values. Its digital representations, like sRGB, therefore, are also defined by integer values of red, green and blue to represent every available colour. However, the gamma operation used to transform between both spaces produces real numbers. By convention, in digital environments, the ideal mathematical RGB space, also called linear RGB due its linear relationship with XYZ, is represented by the possible range of values between 0 and 1. Since XYZ is a space of real values, allowing the intermediate linear RGB to feature real too, error by rounding before applying the gamma is avoided. In order to know the true RGB value of the according linear RGB and sRGB colours, it is only needed to multiply the linear RGB coordinates by the maximum channel value 2^n, n being the number of bits/pixel in usage.

The linear RGB to XYZ transformation is achieved by a matrix multiplication, referenced in equation (5). However, the matrix coefficients depend on the digital variant of RGB the image is represented with. The recommendation that defines the digital RGB in usage indicates which chromaticity coordinates have to be chosen for calculating the primaries and reference white in the chromaticity plane. Thus, XYZ is defined specifically for a given RGB variant knowing its primary components and its maximum possible value.

The following equations (5) and (6) define the process from transforming between linear RGB and XYZ [51]. The intermediate square matrix allows the transformation from an RGB point in space, defined by its R, G and B values to its counterpart point defined by X, Y and Z. X_w, Y_w and Z_w define the XYZ coordinates for the reference white point, which usually will be the one defined by a standardized illuminant. The intermediate parameters $S_r, S_g, S_b, X_r, Z_r, X_g, Z_g, X_b$ and Z_b are calculated using the expressions (7) and (8) with the chromaticity coordinates x_r, y_r, x_g, y_g, x_b and y_b for the primaries. Y_r, Y_g and Y_b are considered 1 in this transformation.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} S_r X_r & S_g X_g & S_b X_b \\ S_r Y_r & S_g Y_g & S_b Y_b \\ S_r Z_r & S_g Z_g & S_b Z_b \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (5)$$

TABLE 1. Table depicting the primary xy coordinates for different RGB digitalisations.

List of primary chromaticity coordinates for different RGB digitalisations								
Space	R primary		G primary		B primary		Reference white	
	x	y	x	y	x	y	x	y
sRGB	0.64	0.33	0.3	0.6	0.15	0.06	0.3127	0.329
DCI-P3	0.68	0.32	0.265	0.69	0.15	0.06	0.314	0.351
Display P3	0.68	0.32	0.265	0.69	0.15	0.06	0.3127	0.329
Rec. BT. 2020	0.708	0.292	0.17	0.797	0.131	0.046	0.3127	0.329
Rec. BT. 2100	0.708	0.292	0.17	0.797	0.131	0.046	0.3127	0.329

$$\begin{pmatrix} S_r \\ S_g \\ S_b \end{pmatrix} = \begin{pmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{pmatrix}^{-1} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} \tag{6}$$

Likewise, in order to obtain the parametric XYZ coordinates for any reference white point X_w, Y_w, Z_w, when one of the CIE standardized illuminants is considered, the process starts by assigning $Y_w = 100.0$. Then, the corresponding X and Z are also calculated employing the equations (7) and (8) and multiplying them by 100. Here, the chromaticity coordinates for the chosen white point x_w and y_w are employed.

$$X_{point} = x_{point}/y_{point} \tag{7}$$

$$Z_{point} = (1 - x_{point} - y_{point})/y_{point} \tag{8}$$

The equations (9), (10) and (11) below represent the particular transformation case for sRGB after having calculated the parameters from above's guideline [51], using D65 2° as the standardized illuminant [44]. Aside, it is possible only calculating the luminance (Y) of a certain RGB point by performing the second operation alone, so information about lightness is easily available, too.

$$X = 0.4124564R + 0.3575761G + 0.1804375B \tag{9}$$

$$Y = 0.2126729R + 0.7151522G + 0.0721750B \tag{10}$$

$$Z = 0.0193339R + 0.1191920G + 0.9503041B \tag{11}$$

In order to reverse the transformation, it is performed by multiplying the XYZ point with the inverse transformation matrix.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} S_rX_r & S_gX_g & S_bX_b \\ S_rY_r & S_gY_g & S_bY_b \\ S_rZ_r & S_gZ_g & S_bZ_b \end{pmatrix}^{-1} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{12}$$

The last step is obtaining CIELAB coordinates. The mathematical theory that grounds the transformation between CIELAB and device-dependent colour spaces such as RGB is highly non-lineal. Furthermore, the necessity of value normalization and scale difference between the original and transformed space results in croppings and operation exceptions for the values in the low extreme of the analytic domain. This is represented by the different equations that define CIELAB depending on the stimuli values and its limitations in space [35]. The main possible means for converting information from the RGB space to CIELAB is by following the formulae that build up the base of CIELAB out of XYZ starting from RGB, since it depicts absolute values, much like XYZ [38], [52]. Thus, a non-uniform and device-dependent space is transformed first into an independent space, and then into a uniform and independent space [41]. The equations (13), (14) and (15) describe the process.

$$L^* = 116f\left(\frac{Y}{Y_0}\right) - 16 \tag{13}$$

$$a^* = 500\left[f\left(\frac{X}{X_0}\right) - f\left(\frac{Y}{Y_0}\right)\right] \tag{14}$$

$$b^* = 200\left[f\left(\frac{Y}{Y_0}\right) - f\left(\frac{Z}{Z_0}\right)\right] \tag{15}$$

The term $f(q)$ with $q = \frac{X}{X_0}, \frac{Y}{Y_0}, \frac{Z}{Z_0}$, as stated in conditions (16) and (17), describes the possible variations in the equations of L, a^* and b^* depending on the value acquired by q, which stands for the relationships between the XYZ coordinates and their normalized values under a perfect diffuser, X_0, Y_0, Z_0, while being in the same illuminant conditions. Therefore, the CIELAB formulae respond differently depending on the position of the considered values in space (especially for low values), and its proportion towards the reference stimuli [35].

$$f(q) = \sqrt[3]{q}; \quad \forall q > 0.008856 \tag{16}$$

$$f(q) = 7.787q + \frac{16}{116}; \quad \forall q \leq 0.008856 \tag{17}$$

The reverse transformation, from CIELAB to XYZ, is obtained by inverting the previous operations, as in (18), (19) and (20).

$$X = g((L^* + 16)/116 + a^*/500) \tag{18}$$

$$Y = g((L^* + 16)/116) \tag{19}$$

$$Z = g((L^* + 16)/116 - b^*/200) \tag{20}$$

$$g(q) = q^3 \quad \forall q > 6/29$$

$$g(q) = 3(\frac{6}{29})^2(q - \frac{4}{29}) \quad else$$

As some last comments on the transformation procedure, there are some important matters to consider, especially when operating in a digital-only world, like in this presented work. XYZ's condition of being device independent and absolute makes it good to be a space for storage and transmission, albeit the need for exactly a certain combination of real numbers to be translated into a particular colour can be a

drawback when a broad rounding is applied on them. As long as the concrete decimal value of XYZ is maintained, biyectivity between linear RGB, and then sRGB too, and XYZ is conserved. Having not only a device-independent, but also a uniform space for colour information, can solve the drawback XYZ is risked to. Due to its definition and its similar relationships between colours, CIELAB allows some rounding of real values to their direct integers without corrupting the digital RGB value on the other end of the transformation process and still describing absolute colours. XYZ as the intermediate step is not affected. On the other hand, since CIELAB does not admit some colours it is important to note that any modification performed in the information has to be in the limits of the admitted range so biyectivity between spaces is respected. Especially to be considered are the vertices of the space, the RGB and YCM primaries. Their La*b* coordinates are real values, so rounding them to the integer is not recommended because the obtained value may fall out of the admissible range, and thus correspond to a radically different perceptual colour. The fact that there exist perceivable RGB combinations that aren't contained in CIELAB's domain and get rounded to their nearest existing value implies that it must be assured, for biyectivity to be fulfilled, to employ a set of colours with an unique correspondent in both ends, RGB and La*b*.

For this work we aim to recover the final L, a* and b* coordinates of the homogenised pictures in RGB for assessing their further evolution in time. Therefore, it is necessary to consider the outlined transformations and theories for achieving a consistent, lossless and coherent transformation of the information in a digital context where biyectivity is assured.

IV. EXPERIMENTAL SETTING

A. METHODS

It has been made clear that this paper aims to present an improvement for the acquisition of crowdsourced images, making sure to stay within the practical limitations of mathematical colour theory and technology so it can be installed in any case. Comparable recent works [10] employed images directly acquired from the cellphone cameras in use, with the addition of a white balancing process using the Fiji open-source Image Processing Package, and evaluated their suitability for providing useful information for the maintenance task by observing the colour difference between the whole set and the colour reference they had, measuring the information in CIELAB via a spectrometer X-Rite 518. Our experiments incorporate the addition of a more refined intermediate step between the acquisition and the digital quality evaluation fully developed by the team behind this work. We also perform all the steps in a digital environment and all measures are digitally extracted, so there exists no need for other physical equipment such as spectrometers. We perform this as a set of operations that engulf switching from one colour space to others and projecting all possible data combinations in an image into the new defined space that is common

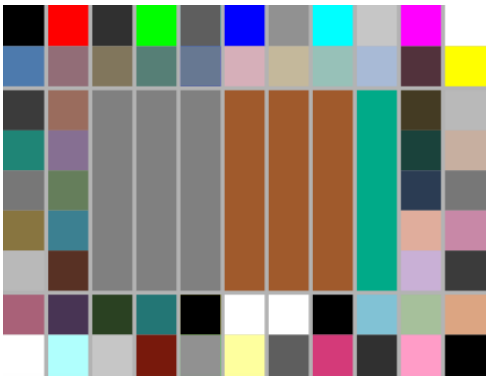


FIGURE 4. Our proposal for the employed colour checkerboard. The squares depict colours that are regularly spread in space, allowing a general calibration without incurring in possible excessive interpolation losses in localized regions of space. The seven rectangular parts in the middle correspond to placeholders where the material tags will be placed in real-life scenarios.

for every picture and close enough to the reference data via an adaptive transfer function calculated by interpolation. For this, we define some fixed points corresponding to certain colours in our reference colour space, and project the rest by performing a linear interpolation between them, thus being able to recover any colour value in the considered space after the calibration. It is an adaptive operation that assures a good improvement in the quality of the results in previous approaches [10] and leaves place for further refinements and even more improvement by proposing alternate and additional operations (figure 5).

With the target in mind of approximating this experimental scenario the most to the possible real-life situation, an controlled environment where visitors may take pictures of the exposed heritage pieces, the following circumstances have been enforced to be fulfilled during the course of our experiments.

In practical cases, as they are intended to be, a necessary condition is the complete control of lighting. Although the availability of ideal luminaries and a totally uniform shining would not be entirely possible, it's adequate to have control over the light source and angle in order to take it into account for the implementation. Lighting modeling is not discarded for fixing possible errors for future steps.

Aside, it is imperative to have a set of colour values of reference in the acquisition that will serve as anchor points for building the transfer function. In other words, pictures need to depict a number of uniform colour samples that define representative points in the reference colour space so the projection and its parameters can be calculated. This issue is solved by establishing the presence of a checkerboard with a selected set of colours in the scenario so it can be included in each image to be taken, as seen in figure 4.

For the outlined and effectuated experiments, pictures depicting the checkerboard only have been taken in controlled lighting conditions. This helps homogenising the colour information of the matter and thus extracting a trustworthy

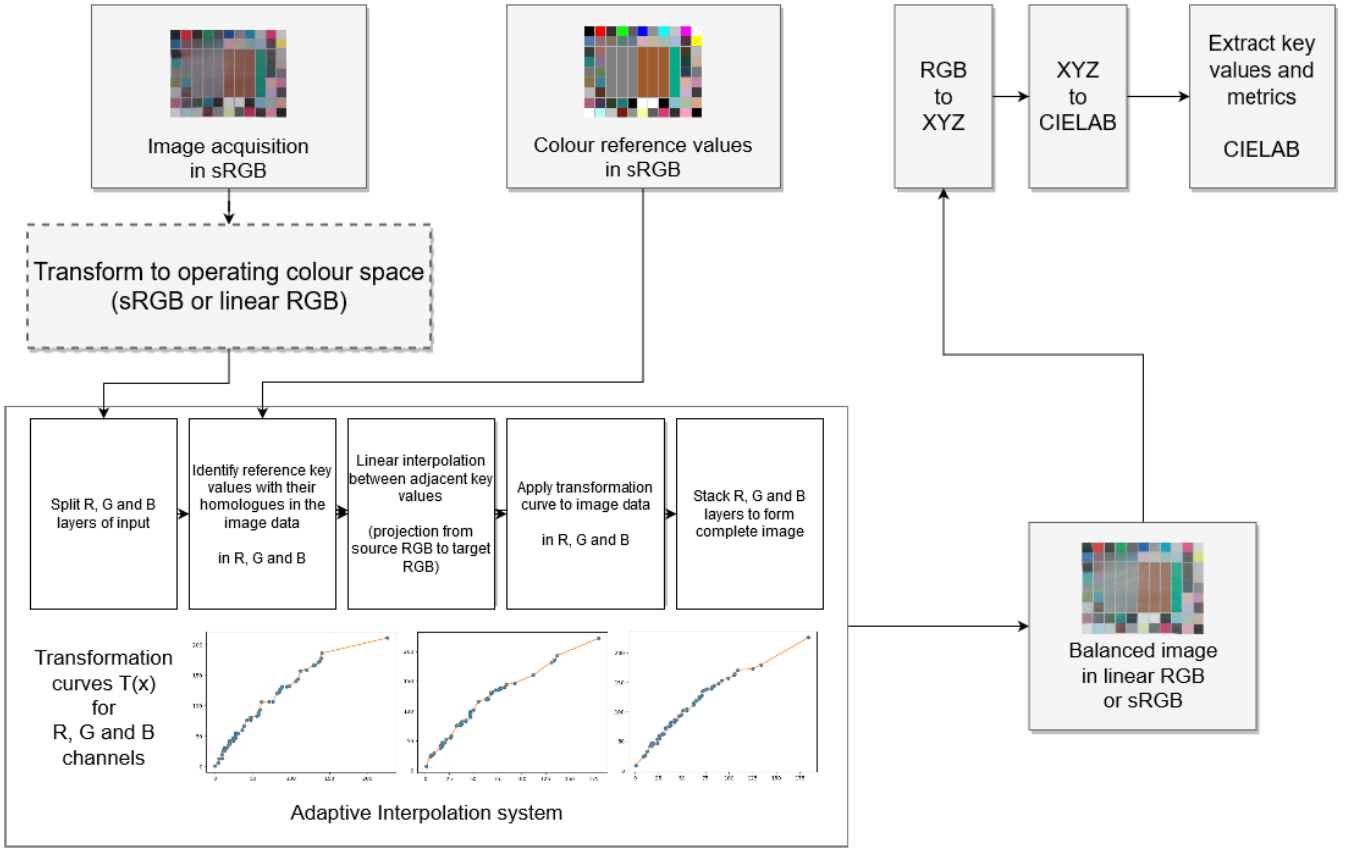


FIGURE 5. **The methodology followed.**

CIELAB information. However, in practical cases, metallic or chemical tags, which will change the colour of their surfaces as indicators of degrading ambient conditions for the heritage pieces [13], will be placed alongside the checkerboard. Even in closed and supposedly controlled environments like museums non-favourable conditions that change the properties of matter can take place, like excessive humidity [20] or the presence of oxidation agents [14]. For future real life scenarios of our proposal, this checkerboard plus the corresponding indicating tags will be put next to the monitored piece, so the atmosphere is shared by both items. The colour information of the indicating tags will be the one employed for the proper surveillance of their corresponding piece under preservation.

Our checkerboard consists of 60 colour patches of dimensions 1 cm $\times$ 1 cm arranged on a chart with 11 columns and 9 rows. In the middle there is reserved a free space of 7 columns $\times$ 5 rows. Here is where the indicator tags will be placed in the future scenarios so they are included in the taken pictures alongside the reference colours.

The employed colours in the checkerboard include regularly sparse representatives of the gray range in RGB and the L coordinate, including maximum black and white, and the primaries red, green, blue, cyan, magenta and yellow. The rest of colours are designed to exist in the perceptual spaces and inside CIELAB's borders, in order to avoid transformation errors, and are combinations of equal divisions of each RGB axis. This way it is assured that the possible range of values is covered equally and inferring the projection of colours not present within the set of anchor colours will not lose much accuracy. Here resides the main difference with checkerboards of standardized usage as the MacBeth type [53]. Whilst those rely on showing a variety of hues which are likely to be found in colour reproduction, such as the primaries or skin colour, ours aims to cover the mathematical space as regularly as possible for accuracy reasons, as a mean for a general calibration.

Our checkerboard is specifically designed to encompass the possible data range in 8-bit RGB and sRGB, covering the whole range without the clippings the gamut followed by UIT-R BT.709 imposes.

We have built a *camera obscura* (figure 6) where the lightning conditions for the photography process can be controlled. The enclosed space consists in a box covered with an uniform black material to avoid undesired reflections with an

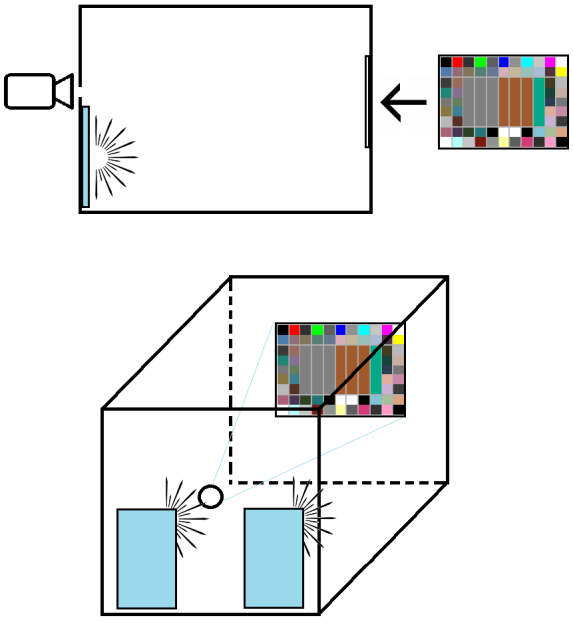


FIGURE 6. Top: Our *camera obscura* as seen sideways. Bottom: Schematic 3D view. The checkerboard lies on the opposite side of the hole and the camera, whilst the lighting comes from two sources below the hole.

only small opening in the middle of one of its sides, where the camera is to be placed. In the same side there stand two lamps INFAIMON BIBL-W100/80. On the opposed side of the box, the reference checkerboard printed on a standard sheet of paper is fixed.

Proceeding like this, an environment notoriously free of lighting aberrations can be obtained, yet not of perfect conditions. This allows proceeding with a set of experiments that serves as an appropriate approach to the intended a practical implementation. Lighting is not ideal but it is controlled, like in the planned future environments to be exploited, and the uniform conditions of the scenario let experimenting with the homogenisation procedure without needing to apply further previous lighting corrections.

The set of images used here is obtained by means of three cameras of three different mobile phones that can be found under public domain; this allows another degree of emulating the conditions of a practical case: obtaining the colour cues from cellphone cameras any visitor could have with them during the visit to the piece. The models of the phones are Samsung Galaxy J30, Xiaomi Redmi 4 and OnePlus 2 (figure 7).

Each camera takes 6 pictures in a row of the checkerboard in the uniform environment of the box. Thus, not only the variability between different mobiles can be treated, but also a possible variability amongst pictures of the same phone can be observed, too. This process is done twice: once with the luminaries working at full intensity (16/16), and the other at half intensity (8/16).

Having installed and conditioned the *camera obscura* in the described way and having the cellphones ready, the image acquisition is performed. The 6 photographs per phone and light conditions are taken while maintaining the field and angle, which is fulfilled by taking the pictures through the hole.

It's important to note that this allows emulating the final requirement that makes the experiments outlined in this paper analogous to the installation of our approach in real life scenarios. When collecting data from the third parties, they would have been previously instructed to shoot the pictures respecting a certain distance and angle with respect to the indicator tags and checkerboard, by fitting a frame to the screens of their phones. This is conceived in order to get a homogeneous flow of information, where outliers with false or reduced information are likely to be infrequent.

After the acquisition, the designed process of calculating the transfer function for the calibration is performed on every component of the set images, which is digitally performed by

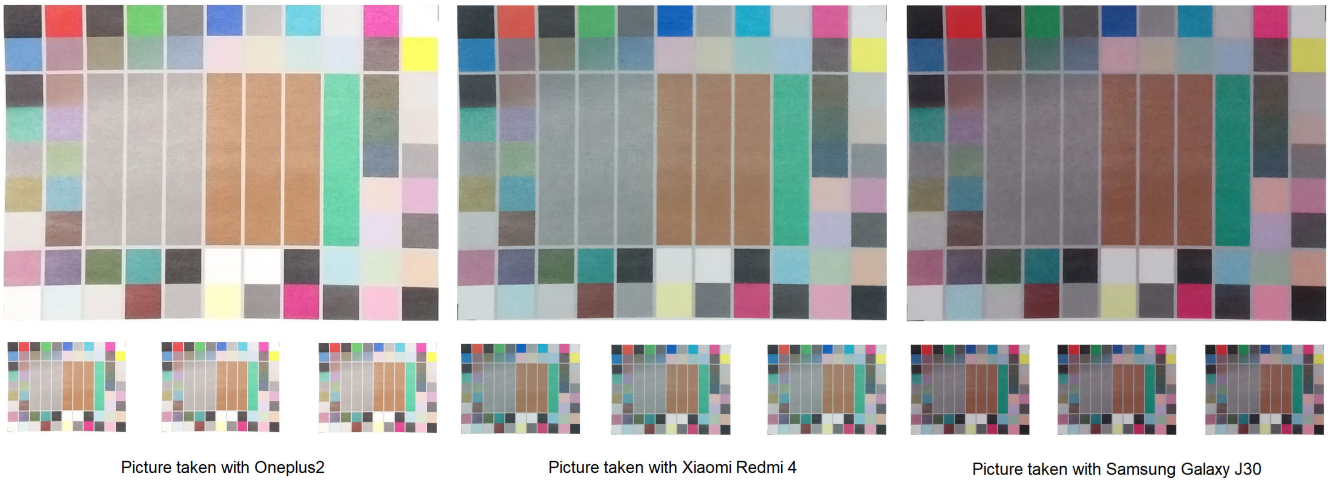


FIGURE 7. Examples of pictures taken with each camera. Top row depicts magnified examples for detailed observation. Bottom row shows the small intra-camera variability for each camera set.

means of Python 3 scripts. The described enforced conditions in the scenario of this experimental process are sufficiently similar to the intended real-life environments to prove that its successful results are validated to work in soon-to-be future installations.

B. EXPERIMENTAL PROCESS

At first, we can observe a first clue of the variability between perceiving cameras, even with a probe set: the images taken by each one, even having captured the same colour and light information in the same circumstances, look different in terms of overall luminosity and hue. From a subjective point of view by any observer, the three sets could be ordered from an overall darker to lighter tone sensation, as it can be seen in figure 7.

The different lighting circumstances of the two experiment scenarios do not alter the general sensation when comparing the view between samples of the different cameras. The lighting perception discrepancies remain in both cases: a set is perceived as the brightest, one as the darkest, and the other lying in between.

After the acquisition is successfully fulfilled, the pictures are stored for their further post-processing and their homogenisation. The next steps of the process are digitally performed employing source code files programmed in Python 3 and image processing and analytical libraries like OpenCV, scikit-learn or pandas.

The post-process mainly consists in operations that cover the whole possible domain for sRGB values, considering we assume we work with images in 8 bits due to the prevalence of 8-bit sRGB data. Each image is separated into its three R, G and B layers, and after their processing they're recomposed for the result analysis, as hinted in figure 5.

First, the exact RGB values for each colour patch in the employed reference are known in beforehand, so they are previously stored within the 60 anchor values $[r_0, r_1 \dots r_{60}]$. The main operation consists in calculating a projection from all the RGB pixel values in each image into the values of the reference picture and building a transfer function $T(x)$ out of it. This is achieved by associating every anchor value with their homologous of the checkerboard in the image to be transformed $[x_0, x_1 \dots x_{60}]$, so the rest of the possible RGB range can be calculated via a linear interpolation between every pair of successive points x_i, x_{i+1} and r_i, r_{i+1}, $i \in [0, 60]$. This approach is simple but enough considering that the reference values r_i are regularly spread throughout the sRGB range, so there is not an extremely big distance between each one of them. Thus, the slope of the transfer function between each successive pair implies a trade-off between low errors and amount of anchor values.

$$T(x) = \begin{cases} \dfrac{r_1 - r_0}{x_1 - x_0}(x - x_0) + r_0 & x \in [x_0, x_1] \\ \dots \\ \dfrac{r_n - r_{n-1}}{x_n - x_{n-1}}(x - x_{n-1}) + r_{n-1} & x \in [x_{n-1}, x_n] \end{cases} \tag{21}$$

So, a projection function $T(x)$ composed by linear slopes between successive anchor points and their counterparts for each single calibrating image is obtained (21), with $n = 60$ being the amount of colour squares. For every image to be transformed there will be three projection functions $T_R(x), T_G(x), T_B(x), x \in [0, 255]$. Following the analysis of the linear interpolations, the expression in (22) states how to obtain any projected value for a given point (y_p) when knowing the pair of points delimiting the slope (x_1, y_1) and (x_2, y_2) and its value to be projected x_p.

Obviously, being an adaptive process, the slope and position of the anchor points may vary on each case, too. The application of this transfer function on the layers of every pixel from the original RGB obtains the calibrated value for the pixels of the final corrected image. It's only necessary to stack the calibrated R, G and B layers to obtain the final result. An example of the transfer curves obtained are depicted in figure 8.

$$y_p = y_1 + (x_p - x_1)\frac{y_2 - y_1}{x_2 - x_1} \tag{22}$$

In the experiment, due to the non-ideal nature of the colour patches in the employed checkerboards, because of them being printed on paper, and the camera acquisition, the digitally uniformly conceived patches are perceived, when printed, as a dispersion of pixels of similar but not equal value instead as a uniform region of a single pure value. This could lead to problems to extract the correct anchor value for every patch. In order to approximate the performed experiment to the most real case, the reference picture employed for building the set of anchor points here is one of the pictures taken in the controlled scenario, namely a picture taken by the Xiaomi Redmi 4 phone (figure 9). This way the colour dispersion on each patch is maintained in the reference and in the gathered images and the tones are assured to lie on a realistic space, so a less extreme compensation has to be performed. The stored anchor values will be then either means or modes of the printed patches. This allows a less complex interpolation calculation and projections to be achieved. Aside, it permits avoiding the possible rupture of biyectivity by associating many differently perceived pixels of every patch to a unique uniform ideal value.

Since the taken pictures depict a controlled environment, any of them is suitable to be chosen as the reference. In this case, we chose one of the images of the intermediate luminosity impression, so compensations made on the other two sets are less complex to achieve. It needs to be said that for this scenario only the images from the other two, non-reference cameras, are compensated.

Under practical effects, the outlined operation works as a post-calibration for the images. The inter-camera variation of the obtained visual information is greatly overcome by performing this process, since the effects and characteristics of the transfer function will be unique and adequate to each picture. Thus, it is suitable for colour surveillance without relying on a single camera if needed. Aside, the possible intra-camera variation defects can be surpassed, too.

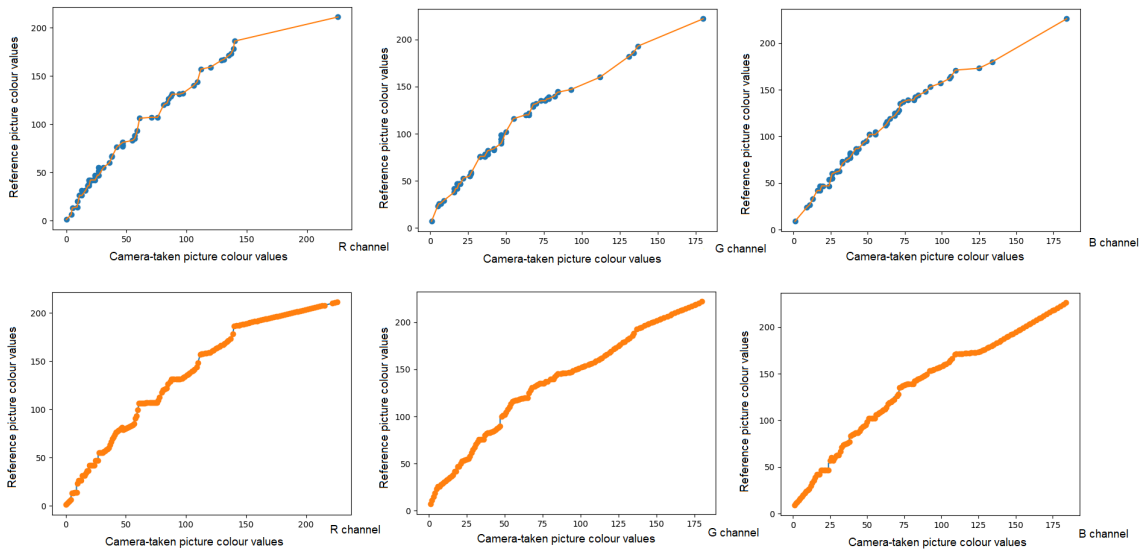


FIGURE 8. Examples of interpolation curves for each image layer in the case of the Samsung Galaxy J30 correction. X-axis represents the domain of the pictures; Y-axis the reference domain. Top row: reference values and interpolation for red, green and blue channels. Bottom row: assigned values for the whole domain for red, green and blue channels.

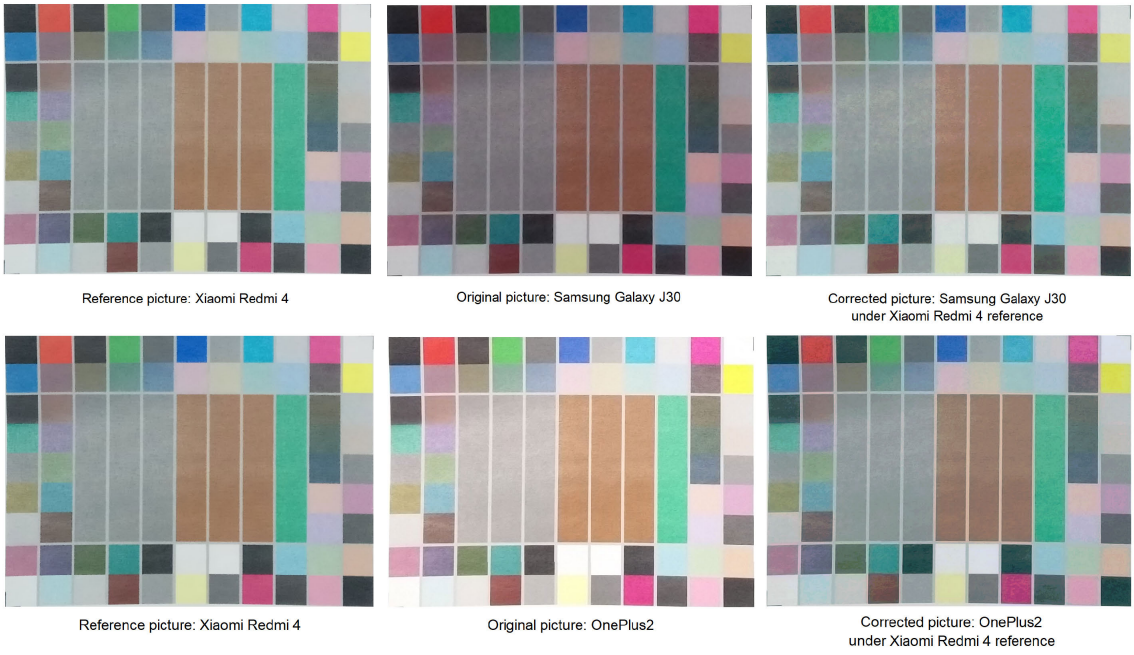


FIGURE 9. Results from the experiments for both cases: Samsung Galaxy J30 and OnePlus2 corrected using Xiaomi Redmi 4 as real reference. The corrections for every set of pictures taken by each camera are analogous given the small intra-camera variability. From left to right: reference picture, original before correction, and original after correction.

In this paper, this compensation operation is performed twice for each image to compare different combinations: once calculated for the digital sRGB values depicted in the images and other for the linear, mathematical RGB values of each patch. This means the necessity of other additional previous transformation operations. If the compensation is to be performed in sRGB, the image data should remain untouched; and if the compensation is designed to be done in linear RGB, the image data need to be transformed from sRGB to linear RGB.

After the whole process, each of the resulting images is analysed by means of different objective metrics. The difference between the reference colour patches and the corrected ones is quantified, alongside the standard colour difference following the CIEDE2000 ΔE_{00} reference formula [54], as seen in equation (23), as shown at the bottom of the next page. This formula calculates a distance-like metric taking into account difference values between lightness, chroma and hue, represented by L, C and H, weighted by the S functions and parametric factors k, and also an interactive

term between C and H for improving its effectiveness for the blue range, represented by R_T. Some works [17] defend the usage of the CIEDE76 difference for heritage science instead, being it more suitable for evaluating colour changes in the proper matter such as stone. However, in this work we are measuring the difference between colour patches of the reference checkerboard in CIELAB instead, regardless of the nature of the future indicator tags. Correlations and means, modes and standard variations and histogram representations are calculated for each patch in order to assess the dispersion for each colour. This is calculated in RGB/sRGB and CIELAB for each colour patch.

When calculating the CIEDE2000 differences between colours, patches of the same size for every colour in the reference and compensated images are employed, and their mode is taken as the representative value for each patch.

Other valid objective quality measure employed is peak signal-to-noise ratio (PSNR), which is taken into account in state-of-art proposals [21]. We establish a quality comparison between patches of the reference image and the same patches for the corrected ones, to assess that not only the final retrievable CIELAB values are satisfactory, but that the whole calibration process for all the possible pixel values is successful and valid for its application on real scenarios.

In the end, we perform an experiment decomposed in scenarios, whose algorithm is here outlined (algorithm 1). We have the two different lighting conditions. Each scenario is a combination of performing the compensations in sRGB and linear RGB and employing as a reference the controlled real values of a photograph of one of the sets. In order to overcome the disparity between cameras and obtain admissible, state-of-art results in the field of crowdsourcing information for heritage science, this variety of experiments addresses the logical first approaches to solve the regarded problems. The results will be discussed in the following section.

Algorithm 1 Algorithm of the Outlined Approach

```
Data: Image acquired with cellphone; checkerboard
      reference values
Result: CIELAB color values
Initialization: acquisition of image;
while Acquired image is unbalanced do
    if Process in linear RGB then
        Transform image data to linear RGB;
        Transform reference data to linear RGB
    else
        go to next section;
    end if
    Split data in R, G and B layers;
    Identify key values of the image with reference;
    Projection curve: linear interpolation between
    adjacent key values;
    Apply projection to all the image data;
    Stack R, G and B to form balanced image;
end while
Transform image to XYZ;
Transform image to CIELAB;
Extract key values and quality metrics;
```

V. RESULTS AND DISCUSSION

After having performed the set of mentioned experiments and adjustments, a resulting feedback consisting of the calibrated pictures and a set of metrics and graphics is obtained.

When visually inspecting the received images, it can be recognized that the overall ''tone'' has been modified (figure 9). The generally reigning visual difference between pictures of different cameras, seen as a lighter or darker tone, is subverted in favour of a common look. Under a visual subjective judgement, the tones have been shifted to match with the given reference in each case, while respecting the slight variance of tone for each patch of colour that is natural for a photograph. It can be said that this process does not trim or clip the information at assigning a new single value to a variety of original ones; the projection assigns different pixel values to every different pixel value in the compensation, as it was conceived when employing as the reference one real photograph. This transfer happens almost identically with each image of the given sets since each projection is calculated effectively for every sample; the possible variations of the outcome may lie on outlying colour values or saturation. Visually, a calibration-like result is achieved and an overall homogenisation of different visual aspects is reached. White and black tones are saturated and the differences between compensated patches of different images origin from pixel arrangements or outliers.

It is also interesting to see that the two different considered lighting conditions for the experiment, one at full intensity, the other at half the full intensity, do not seem to be significant when performing the projection. When inspecting the original pictures, while it is true that the difference in intensity is noticeable in the different sets, the adaptive nature of our processing fixes this possible disparity as long as the light ambiance is not aberrant to the point of introducing notorious errors.

In order to objectively evaluate the quality of the results given by the experiments, statistical distribution values and measures of correlation and difference have been taken. First of all, histograms for a square area comprising the center for each layer for each colour patch and a surrounding

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{k_C S_C}\right)\left(\frac{\Delta H'}{k_H S_H}\right)} \tag{23}$$

TABLE 2. Table depicting the correlation measures by layer using representative colour values for the whole image between reference and balanced samples.

Value	Half light, sRGB	Half light, RGB	full light, sRGB	full light, RGB
Pearson for L layer using modes	0.9949	0.9952	0.9949	0.9841
Pearson for a* layer using modes	0.9851	0.9833	0.9890	0.9718
Pearson for b* layer using modes	0.9806	0.9787	0.9787	0.9978
Pearson for L layer using means	0.9981	0.9979	0.9983	0.9914
Pearson for a* layer using means	0.9896	0.9855	0.9936	0.9936
Pearson for b* layer using means	0.9856	0.9857	0.9848	0.9793

TABLE 3. Examples of average PSNR measures.

Sample	PSNR with sRGB balance	PSNR with RGB balance
Picture 1	36.748	33.335
Picture 2	36.633	33.040
Picture 3	36.775	33.372
Picture 4	36.835	33.422
Picture 5	36.672	33.114
Picture 6	36.739	33.333
Picture 7	39.419	36.166
Picture 8	39.260	35.414
Picture 9	39.484	35.551
Picture 10	39.331	35.930

of 35 pixels are calculated, as well as their mean, mode and standard deviation. It is desirable for the mean and the mode to be reasonably similar so all the pixel values taken effectively belong to the same patch. Then, these measures are compared between pictures of the same camera (intra-camera) and between different cameras (inter-camera). It can be mathematically seen that consistency and stability in the samples has been greatly achieved. The intra-camera variability between samples has been overcome when noticing the extremely similar values of the histogram statistics, and the homogenisation noticed in the visual evaluation is confirmed by seeing that the only relevant differences lie within the pixels resulting as outlying values during the interpolation. Overall, the intra-class variability is more accused on standard deviation and modes values whilst means stay almost identical. The effects and results are the same regardless on which space the correction is performed (RGB or sRGB). The final concrete pixel values vary in each scenario though giving a brighter or darker tone to the images, due to the gamma operation being performed before the compensation or after it depending on the case, but the observed properties stay. Nevertheless, the compensation performed in sRGB delivers images more similar to the original reference.

Aside, an error metric in form of the CIE colorimetric distance CIEDE 2000 is taken. This formula [54] calculates an effective weighted mathematical distance between two samples specifically for the CIELAB information, since it considers its value for multiple dimensions. Being a difference value, it is perfect for evaluating the closeness and possible error between the same colour in two different images objectively. In the performed experiments we employ the formula to compare the homogenised results with the references, and thus evaluate the precision of the compensation. Since the CIELAB coordinates are our final target to get, we compare the LAB coordinates with the distance metric. The magnitude of the calculated distances for each colour is usually in the range between 0.2 and 3.5, with very extreme isolated cases reaching higher numbers in presence of outlying values.

In order to establish a valid comparison with the work of [10], we also calculated the average difference between each colour patch between all the balanced samples and the reference. While their results consist in difference values mostly superior to 5 and most of them even higher than 10, our results, which can be verified in table (4), prove that our added image processing technique for correction and balance improve the quality of the information for their further usage. From our analysed 60 patches, only 6 of them bear a difference superior to 5, and five of those are between 5 and 7. The range of the rest of differences lies between 0 and 3.5, and a fewer amount higher than 4. The tolerance, seen as standard deviation from the mean differences, lies usually between 0 and 1, and the extreme cases reaching values superior to 3 are rarely found. A distance of 0.0 is the perfect similitude, and the higher the magnitude the more different the samples are. Display technology manufacturers state that the minimal detectable difference lies between the values of 1 and 2.5. When establishing a comparison to assess the validity of our results, it can be seen that, typically, for commercial

TABLE 4. Statistics of average differences taking every picture and colour patch into account.

	Balance in sRGB				Balance in RGB			
Patch	Diff.	Dev.	Dif. %	Dev. %	Diff.	Dev.	Diff. %	Dev. %
0	3.774	1.082	2.771	0.794	4.855	1.102	3.565	0.809
1	2.456	0.636	1.803	0.467	1.928	2.266	1.415	1.664
2	2.565	1.728	1.883	1.269	3.094	0.517	2.271	0.38
3	1.358	0.812	0.997	0.596	2.951	2.033	2.166	1.493
4	0.728	0.476	0.534	0.35	0.873	0.607	0.641	0.446
5	1.729	0.652	1.269	0.478	2.308	0.876	1.694	0.643
6	2.8	1.145	2.055	0.841	3.905	2.075	2.867	1.523
7	2.069	0.961	1.519	0.705	3.059	1.422	2.246	1.044
8	2.243	0.975	1.647	0.716	5.117	3.586	3.757	2.633
9	11.734	1.705	8.615	1.252	15.147	3.739	11.121	2.745
10	5.926	1.407	4.351	1.033	10.723	3.611	7.873	2.651
11	2.633	0.86	1.933	0.632	3.468	1.022	2.546	0.75
12	1.646	0.364	1.209	0.267	2.676	0.695	1.965	0.511
13	2.589	0.148	1.901	0.109	3.389	0.47	2.488	0.345
14	2.925	0.271	2.147	0.199	3.551	0.504	2.607	0.37
15	2.188	0.359	1.607	0.264	2.775	0.861	2.037	0.632
16	2.473	0.678	1.815	0.498	3.122	1.052	2.292	0.772
17	1.461	1.294	1.072	0.95	2.782	0.526	2.043	0.386
18	2.302	0.438	1.69	0.322	2.135	0.39	1.568	0.286
19	3.147	1.115	2.31	0.819	3.755	1.118	2.757	0.821
20	3.982	1.189	2.924	0.873	5.589	2.477	4.104	1.818
21	1.448	0.524	1.063	0.385	1.925	0.966	1.413	0.709
22	1.993	0.319	1.463	0.234	1.668	0.642	1.225	0.472
23	3.815	1.079	2.801	0.793	4.956	2.033	3.639	1.493
24	1.939	1.165	1.423	0.855	3.296	1.798	2.42	1.32
25	2.167	0.959	1.591	0.704	2.417	1.12	1.774	0.822
26	3.352	1.758	2.461	1.29	4.956	2.595	3.639	1.905
27	2.708	1.236	1.989	0.907	3.694	1.432	2.712	1.051
28	0.952	0.235	0.699	0.172	1.604	0.522	1.178	0.383
29	1.844	0.989	1.354	0.726	2.379	0.674	1.746	0.495
30	2.904	1.657	2.132	1.217	3.382	2.036	2.483	1.495
31	2.058	0.345	1.511	0.253	1.899	0.374	1.395	0.275
32	2.134	0.936	1.567	0.687	2.567	0.973	1.885	0.714
33	3.26	2.409	2.393	1.769	4.89	3.705	3.59	2.72
34	1.206	0.733	0.886	0.538	2.161	1.03	1.586	0.756
35	2.735	0.788	2.008	0.578	3.964	1.24	2.91	0.911
36	4.888	1.132	3.588	0.831	6.186	1.376	4.542	1.01
37	3.758	0.134	2.759	0.098	3.326	1.021	2.442	0.75
38	1.451	0.469	1.065	0.344	1.67	0.924	1.226	0.678
39	1.384	0.573	1.016	0.421	1.685	0.913	1.237	0.67
40	2.227	1.596	1.635	1.172	2.966	2.045	2.178	1.501
41	2.295	1.405	1.685	1.032	2.64	1.331	1.938	0.977
42	1.564	0.766	1.148	0.562	2.293	1.396	1.683	1.025
43	0.882	0.61	0.647	0.448	1.764	0.584	1.295	0.428
44	3.01	0.8	2.21	0.588	3.3	1.056	2.423	0.776
45	1.852	1.277	1.36	0.938	1.871	1.672	1.374	1.227
46	1.324	0.323	0.972	0.237	1.543	0.637	1.133	0.468
47	0.664	0.204	0.487	0.15	0.622	0.126	0.457	0.093
48	2.076	1.095	1.524	0.804	1.984	0.613	1.457	0.45
49	4.232	1.242	3.107	0.912	6.211	2.573	4.56	1.889
50	3.104	0.82	2.279	0.602	4.091	0.994	3.004	0.73
51	1.644	0.792	1.207	0.581	2.241	0.834	1.646	0.612
52	1.536	0.765	1.128	0.561	1.7	0.801	1.248	0.588
53	2.673	0.957	1.962	0.703	3.787	1.096	2.78	0.805
54	4.112	0.805	3.019	0.591	6.397	1.922	4.696	1.411
55	3.896	1.828	2.86	1.342	6.708	2.078	4.925	1.525
56	3.155	1.154	2.316	0.847	3.529	1.211	2.591	0.889
57	2.269	1.297	1.666	0.952	3.574	0.687	2.624	0.504
58	1.84	1.533	1.351	1.126	2.281	1.667	1.675	1.224
59	3.4	0.422	2.496	0.31	4.139	0.599	3.039	0.439

reproduction a difference between 3 and 6 is considered acceptable [55], so it can be said that our technique offers satisfactory results to the human perception and that it features valid results for its extensive usage for the followed aims.

In the same table we also included not only the CIEDE 2000 difference information in quantitative values, but also as percentage scores, taking the maximum possible difference value between the most distant opposite points of CIELAB $\Delta E_{max} = 136.205$ as 100%. We can see that the vast majority of the values comprises the range between 0% and 4% of the total possible difference.

As a correlation measure we used the Pearson coefficient, which is effectively and widely used for image processing and image validation tasks. Their calculation, using modes and means as representative colour values for all the information in each patch, takes the information of the whole image into account, so it has been made sure during the experiments that the reference and the samples depict the same content without borders and are depicted on the same resolution. It's calculated for each of the layers. The obtained range of Pearson coefficients oscillates between 0.9718 and 0.9983 for all the layers, which means an excellent degree of correlation (table 2).

When calculating the PSNR results, they are performed so the quality of the calibration for each colour patch can be evaluated. Every colour square from the corrected images is compared with their counterpart in the reference. Then, the average is obtained for every image. Table 3 depicts 10 obtained values for the balance in linear RGB and in sRGB scenarios. Pictures 1-6 correspond to the camera with visually brighter results and pictures 7-10 to the visually darker original samples. It can be seen that the values surpass in quality to state-of-art color transfer and correction referential technique [21], up to between 3 and 9 dB more.

Lastly, some visual representations depicting the correlation between original data and the homogenised information are provided to corroborate our results in an intuitive way (figure 10). The averages for each colour patch in every image of the considered set with the standard deviation depending on the pixel value variety of each are depicted in boxplot form in a quadrant. The X-axis represents the range of values in the original space, and the Y-axis the range of the space of projection. The closer the boxplots are to the diagonal, the more accurate the results of the homogenisation are, since it achieves projecting each colour value to its counterpart. The three coordinates L, a* and b* show that a generally good projection is achieved, with outliers lying not excessively far away from the diagonal. The main improvement with respect to the previous work [10] can be observed in the a* coordinate, where the correction shows all values being precisely projected, in comparison to the original sparse boxplots. Our work can add to the original statement that cellphones are a potentially useful sources of measurements in L and b* coordinates not only a better accuracy, but also a more reliable way of providing a* coordinates in this kind of applications;

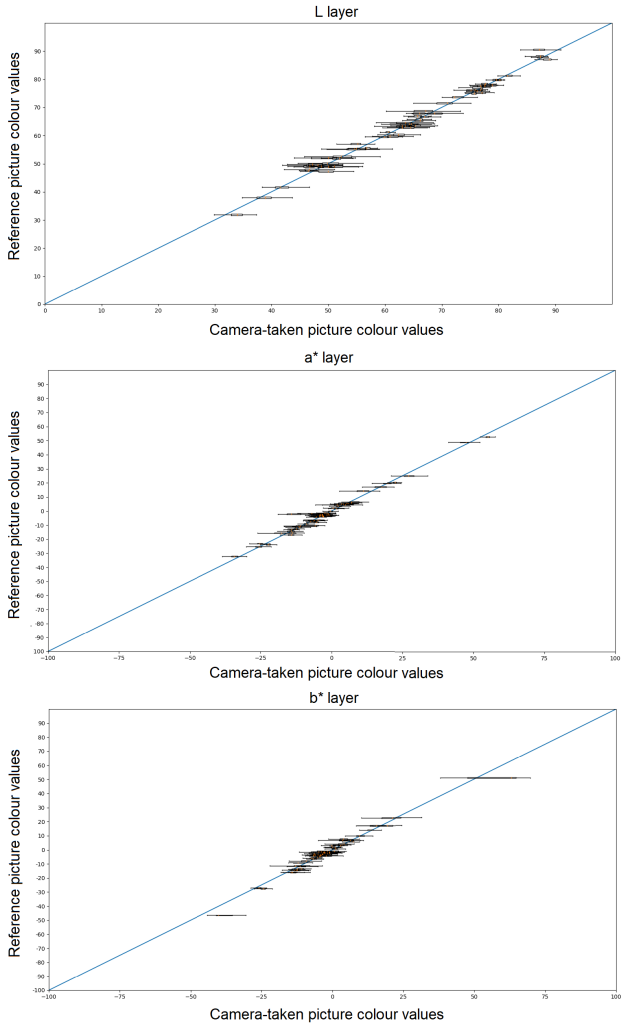


FIGURE 10. Boxplot showing the accuracy of our process, taking all the image data from all the calibrations into account. X-axis represents the domain of the original images; Y-axis the reference domain. From top to bottom, information for all the L, a* and b* data. The nearer to the blue diagonal, the more accurate the results are.

thus proving as an effective and valid solution for heritage maintenance.

VI. CONCLUSIONS AND FUTURE STEPS

Overall, it can be recognized that incorporating our developed image processing techniques for an adaptive white and colour calibration and homogenisation of samples by projection and interpolation, improves greatly the quality of the visual information for our purposes of heritage maintenance in controlled environments compared to previous approaches, judging by the displayed difference [10] and PSNR metrics [21]. Its simpleness also avoids a reinforcement of the mentioned random uncertainty that may happen to chemical colour changes, and lays ground for future refinements and improvements.

The results we have explained in the previous section cover objective measures, according to the nature of our monitoring aims. Correlations are also calculated so the degree

of similarity between reference and correction can be seen. Their analysis reveals that our calibration approach offers an excellent quality of results. The fulfilled experiments are performed in an environment whose conditions are analogous to real settings, so the validity of the results and the overall approach is assessed.

Our work presents the results exclusively in an all-digital environment, with every employed gear conceived for 8-bit perceptual and CIELAB colour spaces, including the checkerboard and all parameters of the performed operations. It engulfs a more concrete work environment than the previously mentioned state-of-art works, that measured mixing spectral measurements with X-Rite and digital information from the pictures [10]. Here we oriented our experiments towards digital operations and the storage of the information provided for its further usage, from the reference values to the ending CIELAB coordinates. Acting like this we assure the usage of the obtained information for assessing an objective measure and observation of its evolution over time in the future system implementation, in correlation with the conditions outlined in section II. The adaptive interpolation will calibrate automatically the colour space and extract the current colour of the indicator tags in the checkerboard.

We also employ a real but optimal piece of information as the colour reference so there is no risk to incur in errors sourcing from trying to reach ideality, like projecting a subrange of values into a single one which could lead to frequency errors, as it has been mentioned before.

While it can seem that operating in an exclusively digital environment may cause some error in the information due to the unavoidable discretization of information, following the colour theory outlined above and avoiding rounding the operations there is no risk of loss. It has to be made sure that XYZ and CIELAB spaces define real numbers, so any rounding may alter the integers in sRGB. Aside, knowing there exist exact borders for the information in CIELAB, it must be made sure that all the colours employed exist within its domain in order to respect the biyectivity of the overall transformation.

The most immediate future step is the practical implementation of the explained calibration technique. The valid results in a near-identical scenario to a controlled preservation environment promise a satisfactory result in real contexts. For this, a colorimetric evolution of materials for the indicator tags will be needed and the extraction of their CIELAB values will be compared with spectrographic measures.

Aside, our current line of scientific research seeks to include transformations directly in the three-dimensional color space that will provide even more accurate adjustments and calibrations than those obtained so far. The nature of those future transformations may extend the basis outlined in this paper on a multidimensional level, or incur in optimizations like Gaussian Mixture Models or regression by artificial neural networks.

Since it could be possible that some person could shoot the picture with a cellphone which uses an RGB space with a gamut wider than sRGB, it will be considered to perform the same set of operations so all the colour data are available in the same fixed space. Switching between spaces and operating in an absolute space will be beneficial in this kind of situations.

Aside, taking note of the accuracy in quality of this approach and the overall non-exppensive possibility of installation, it has to be thought of extending this implementation of heritage preservation of any kind, not only located in concrete controlled environments. Possible changes on the colour distribution in the checkerboard may take place to cover different regions of the colour spaces in usage, in order to assure a greater accuracy in calculating unique colour values present in different materials or lighting conditions, depending on their chemical composition and the way their appearance denotes their degradation under various adverse conditions.

ACKNOWLEDGMENT

The figures depicting colour space solids are created and owned by Mike Horvath [36]. Figure 3 is a slight modification of the licensed file in [42], which states its original owner.

REFERENCES

[1] UNESCO. (2015). *Tangible Cultural Heritage*. [Online]. Available: http://www.unesco.org

[2] R. Nespeca, "Towards a 3D digital model for management and fruition of ducal palace at Urbino. An integrated survey with mobile mapping," *Scires-IT*, vol. 8, pp. 1–14, Jan. 2018.

[3] S. Battiato, G. M. Farinella, F. Milotta, A. Ortis, L. Addesso, A. Casella, V. D'Amico, and G. Torrisi, "The social picture," in *Proc. ACM Int. Conf. Multimedia Retr.*, 2016, pp. 397–400.

[4] K. E. McCracken and J.-Y. Yoon, "Recent approaches for optical smartphone sensing in resource-limited settings: A brief review," *Anal. Methods*, vol. 8, no. 36, pp. 6591–6601, 2016, doi: 10.1039/C6AY01575A.

[5] C. Daffara, G. Marchioro, and D. Ambrosini, "Smartphone diagnostics for cultural heritage," *Proc. SPIE*, vol. 11058, Jul. 2019, Art. no. 110581K.

[6] K. Makantasis, A. Doulamis, N. Doulamis, and M. Ioannides, "In the wild image retrieval and clustering for 3D cultural heritage landmarks reconstruction," *Multimedia Tools Appl.*, vol. 75, no. 7, pp. 3593–3629, Aug. 2014.

[7] M. Khalaf, "Smart cultural heritage: Technologies and applications," in *Proc. 2nd Smart Cities Symp. (SCS)*, 2019, pp. 1–6.

[8] L. Toumanidis, E. Bocaj, P. Kasnesis, and C. Z. Patrikakis, "Supporting cultural heritage preservation through game-based crowd-sourcing," in *Proc. Int. Conf. Strategic Innov. Marketing Tourism (ICSIMAT)*. Attica, Greece: Univ. of West Attica, Oct. 2018.

[9] N. Wang, X. Zhao, L. Wang, and Z. Zou, "Novel system for rapid investigation and damage detection in cultural heritage conservation based on deep learning," *J. Infrastruct. Syst.*, vol. 25, no. 3, Sep. 2019, Art. no. 04019020.

[10] R. Brigham, J. Grau-Bové, A. Rudnicka, M. Cassar, and M. Strlic, "Crowdsourcing as an analytical method: Metrology of smartphone measurements in heritage science," *Angew. Chem.*, vol. 130, no. 25, pp. 7545–7549, Jun. 2018.

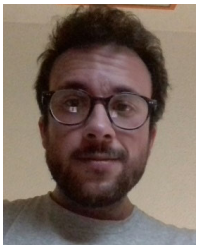
[11] P. Han, D. Dong, X. Zhao, L. Jiao, and Y. Lang, "A smartphone-based soil color sensor: For soil type classification," *Comput. Electron. Agricult.*, vol. 123, pp. 232–241, Apr. 2016.

[12] *Specification for Managing Environmental Conditions for Cultural Collections*, P. British Standards Institution, London, U.K., 2012.

[13] M. K. LaGasse, K. McCormick, Z. Li, H. Khanjian, M. Schilling, and K. S. Suslick, "Colorimetric sensor arrays: Development and application to art conservation," *J. Amer. Inst. Conservation*, vol. 57, no. 3, pp. 127–140, Jul. 2018, doi: 10.1080/01971360.2018.1495480.

[14] W. Ma, M. Walton, O. Cossairt, G. Bearman, and E. Doehne, "Crowdsourced mobile phone images for heritage conservation monitoring," in *Proc. Digit. Heritage*, Sep. 2015, pp. 111–114.

[15] S. M. Goñi and V. O. Salvadori, "Color measurement: Comparison of colorimeter vs. Computer vision system," *J. Food Meas. Characterization*, vol. 11, no. 2, pp. 538–547, Jun. 2017.
[16] N. Fdhal, M. Kyan, D. Androutsos, and A. Sharma, "Color space transformation from rgb to cielab using neural networks," in *Advances in Multimedia Information Processing—PCM*, P. Muneesawang, F. Wu, I. Kumazawa, A. Roeksabutr, M. Liao, and X. Tang, Eds. Berlin, Germany: Springer, 2009, pp. 1011–1017.
[17] A. Molada-Tebar, Á. Marqués-Mateu, and J. L. Lerma, "Correct use of color for cultural heritage documentation," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. IV–2/W6, pp. 107–113, Aug. 2019.
[18] F. L. M. Milotta, G. Furnari, C. Quattrocchi, S. Pasquale, D. Allegra, A. M. Gueli, F. Stanco, and D. Tanasi, "Challenges in automatic munsell color profiling for cultural heritage," *Pattern Recognit. Lett.*, vol. 131, pp. 135–141, Mar. 2020.
[19] T. Wess, "Smartphone citizen science: Can a conservation hypothesis be tested using non specialist technology?" *Heritage Sci.*, vol. 5, no. 1, p. 35, Dec. 2017.
[20] E. Angelini, F. Civita, S. Corbellini, D. Fulginiti, A. Giovagnoli, S. Grassini, and M. Parvis, "Innovative monitoring campaign of the environmental conditions of the stibbert museum in florence," *Appl. Phys. A, Solids Surf.*, vol. 122, no. 2, p. 123, Feb. 2016.
[21] M. Grogan and R. Dahyot, "L_2 divergence for robust colour transfer," *Comput. Vis. Image Understand.*, vol. 181, pp. 39–49, Apr. 2019.
[22] M. Deegan and B. E. Musa, "Preserving the cultural heritage of sudan through digitization: Developing digital sudan," in *Proc. Digit. Heritage Int. Congr. (DigitalHeritage)*, Oct. 2013, pp. 485–487.
[23] EC. (2015). *Europe in a Changing World—Inclusive, Innovative and Reflective Societies*. [Online]. Available: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-societies_en.pdf
[24] J. P. Cohn, "Citizen science: Can volunteers do real research?" *BioScience*, vol. 58, no. 3, pp. 192–197, Mar. 2008, doi: 10.1641/B580303.
[25] M. Haklay, *Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation*. Dordrecht, The Netherlands: Springer, 2013, pp. 105–122, doi: 10.1007/978-94-007-4587-2_7.
[26] C. Nalad, T. Charoenporn, and N. Tongtep, "The monitoring management system for tangible cultural heritage surveillance," in *Proc. 2nd Int. Conf. Adv. Inform., Concepts, Theory Appl. (ICAICTA)*, Aug. 2015, pp. 1–5.
[27] A. Cirafici, V. Amoretti, V. Di Fratta, I. Di Natale, and D. Lallone, "Rethorics through images. New strategies to communicate, valorise and share the cultural heritage," in *Proc. Int. Conf. Virtual Syst. Multimedia (VSMM)*, Dec. 2014, pp. 60–65.
[28] T. Tore and E. Töre, "Documenting 'meaning': A participant model for tangible heritage documentation by social media," in *Proc. Digit. Heritage Int. Congr. (DigitalHeritage)*, Oct. 2013, p. 447.
[29] G. Bearman, E. Doehne, J. Voss, K. Merrill, and R. Ba-Garia, "Citizen science and mobile phone cameras as tools for monitoring world heritage," in *Proc. Built Heritage Monit. Conservation Manage. Conf.* Milan, Italy: ANE Image, Jan. 2013.
[30] R. Johnston-Feller, *Color Science in the Examination of Museum Objects: Nondestructive Procedures* (Tools for Conservation). Los Angeles, CA, USA: Getty Conservation Institute, 2001. [Online]. Available: http://hdl.handle.net/10020/gci_pubs/color_science
[31] B. Fitzner and K. Heinrichs, "Damage diagnosis at stone monuments—Weathering forms, damage categories and damage indices," *Understand. Manag. Stone Decay*, vol. 45, pp. 12–13, Jan. 2001.
[32] V. Lebrun, C. Toussaint, and E. Pirard, "Monitoring color alteration of ornamental flagstones using digital image analysis," Univ. de Liège, Liège, Belgium, Tech. Rep., 2004.
[33] M. Grogan, R. Dahyot, and A. Smolic, "User interaction for image recolouring using $2," in *Proc. 14th Eur. Conf. Vis. Media Prod. (CVMP)*, Dec. 2017, pp. 1–10.
[34] S. Farnand, Y. Jang, L. K. Choi, and C. Han, "A methodology for perceptual image quality assessment of smartphone cameras–color quality," *Electron. Imag.*, vol. 2017, no. 12, pp. 95–99, Jan. 2017.
[35] N. Ohta, "Colorimetry, fundamentals and applications," in *Wiley-IST Series in Imaging Science and Technology*. Hoboken, NJ, USA: Wiley, 2005.
[36] M. Horvath. (2018). [Online]. Available: https://github.com/mjhorvath/Mike-Wikipedia-Illustrations
[37] *Colorimetry*. CIE, International commision on Illumination, Vienna, Austria, 2004.
[38] D. Yanzhe, "A novel algorithm based on improved BP neural network and its application in color management," in *Proc. 7th Int. Conf. Innov. Mobile Internet Services Ubiquitous Comput.*, Jul. 2013, pp. 650–653.
[39] K. León, D. Mery, F. Pedreschi, and J. León, "Color measurement in L*a*b* units from RGB digital images," *Food Res. Int.*, vol. 39, no. 10, pp. 1084–1091, Dec. 2006.
[40] G. Hoffman. (2019). *CIELAB Color Space*. [Online]. Available: http://docs-hoffmann.de/cielab03022003.pdf
[41] M.-S. Cho, B.-H. Kang, and M. R. Luo, "Device calibration of a color image scanner digitizing system by using neural networks," in *Proc. Int. Conf. Neural Netw. ICNN*, Dec. 1995, pp. 59–62.
[42] (2007). *Sakurambo*. [Online]. Available: https://commons.wikimedia.org/wiki/File:CIExy1931.svg
[43] (1999). *CIE Standard Illuminants for Colorimetry*. [Online]. Available: http://www.cie.co.at/publications/colorimetric-illuminants
[44] (1998). *IEC, IEC/4WD 61966-2-1: Default RGB Colour Space—sRGB*. [Online]. Available: https://web.archive.org/web/20050105122556/http://www.colour.org/tc8-05/Docs/colorspace/61966-2-1.pdf
[45] D. Pascale, "A review of RGB color spaces," BabelColor Co., Montreal, QC, Canada, Tech. Rep., 2003.
[46] Technicolor. (2016). *Color Spaces by Technicolor*. [Online]. Available: https://web.archive.org/web/20160203035423/http://www.technicolor.com/en/solutions-services/technology/technology-licensing/technicolor-color-certified/color-certification-process/color-spaces
[47] D. Jackson. (2016). *Improving Color Web*. [Online]. Available: https://webkit.org/blog/6682/improving-color-on-the-web/
[48] *Eg 432-1:2010—SMPTE Engineering Guideline—Digital Source Processing—Color Processing for D-Cinema*, document EG 432-1:2010, SMPTE, Nov. 2010, pp. 1–81.
[49] ITU-Rb. (2015). *Recommendation ITU-R BT.2020-2*. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2020-2-201510-I!!PDF-E.pdf
[50] ITU-R. (2018). *Recommendation ITU-R BT.2100-2*. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2100-2-201807-I!!PDF-E.pdf
[51] B. Lindbloom. (2017). *RGB/XYZ Matrices*. [Online]. Available: http://www.brucelindbloom.com/index.html?Eqn_RGB_XYZ_Matrix.html
[52] C. Connolly and T. Fleiss, "A study of efficiency and accuracy in the transformation from RGB to CIELAB color space," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 1046–1048, Jul. 1997.
[53] IEC. (2011). *CIE 199:2011: Methods for Evaluating Colour Differences in Images*. [Online]. Available: http://cie.co.at/publications/methods-evaluating-colour-differences-images
[54] P. J. Alessi *et al.*, "Technical report CIE 199:2011 methods for evaluating colour differences in images," CIE, Vienna, Austria, Tech. Rep., 2011.
[55] ViewSonic. (2020). *ViewSonic's Delta E $\leq$ 2 Color Accuracy*. [Online]. Available: https://color.viewsonic.com/explore/content/DeltaE%E2%89%A62Color-Accuracy_2.html



MIGUEL ANTONIO BARBERO-ÁLVAREZ received the bachelor's degree from the Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, ETSI Telecomunicación of the Universidad Politécnica de Madrid (UPM), in 2016, where he graduated with a bachelor's thesis about synthesizing contents for UHD television via super-resolution. From 2016 to 2019, he continued his career, coursing the Máster Universitario en Ingeniería de Telecomunicación, UPM, and finished it with a master's thesis about the exploration of machine learning techniques for person re-identification. Meanwhile, from 2018 to 2019, he held an internship for developing machine learning systems in TensorFlow and Keras for emotion recognition and detection of sleep apnea out of speech data in the GAPS Group, UPM. In April 2019, he entered GATV Group as a Predoctoral Researcher, concentrating his work in colorimetric analysis and image processing using algebraic means and deep learning techniques, thus continuing in the same field of his master specialization.

JOSÉ MANUEL MENÉNDEZ (Senior Member, IEEE) is currently the Director of the Visual Telecommunication Application Research Group (GATV), from 2004 to 2019, and the Director of the Chair of the Spanish Public Broadcaster RTVE, Universidad Politécnica de Madrid (UPM), since 2015. He is currently a Full Professor with the Signal, Systems, and Radio-communications Department, E.T.S. Ingenieros de Telecomunicación, UPM, where he has been teaching on topics related to communications, audio-visual systems, and computer vision since 1994. He has extensive experience in participating in and directing research projects (more than 150), both national and European, on topics related to visual communications, digital television, and computer vision. He has published more than 200 articles on these subjects in international journals and conferences, and has been invited to give more than 70 invited lectures (both national and international). He is the author of three granted patents and six software registrations.

Mr. Menéndez is a regular reviewer of the IEEE Signal Processing Society, since 2000, for different international journals and conferences sponsored by that entity. He is also an evaluator of the IET Society for several journals since 2009. He also collaborates regularly with different national and regional entities and the European Commission in the certification of accredited laboratories, in the evaluation and review of R+D+i projects and performs consulting work for broadcasters and companies in the telecommunications sector in Spain. He has participated in the creation of two technology-based spin-off companies, and collaborated in the launch of two other companies in the Information and Communications Technology Sector.

JUAN ANTONIO RODRIGO received the Telecom Engineer degree (Hons.), the master's degree in communications technologies and systems, and the Ph.D. degree *(cum laude)* from the Universidad Politécnica de Madrid (UPM), in 2007, 2010, and 2016, respectively.

He is a member of the Visual Application Telecommunication Group (GATV), since 2007. During these years, his work has been focused on Research and Development projects in a few different areas, including HDTV and 3DTV real-time video processing and coding, video streaming, video quality analysis, computer vision, machine learning, embedded programming for real-time video processing, and industrial process automation. Several international congresses and publications endorse these Research and Development deployments.