

Received September 21, 2021, accepted October 6, 2021, date of publication October 8, 2021, date of current version October 15, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3119223

# Self-Designed Colour Chart and a Multi-Dimensional Calibration Approach for Cultural Heritage Preventive Preservation

MIGUEL ANTONIO BARBERO-ÁLVAREZ<sup>1</sup>, JUAN ANTONIO RODRIGO,  
AND JOSÉ MANUEL MENÉNDEZ<sup>2</sup>, (Senior Member, IEEE)

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 Monitoring by Image Processing and Citizen Science for Conservation of Cultural Heritage Materials Project (Monitorización por procesamiento de imagen y ciencia ciudadana para la conservación de materiales del patrimonio cultural—MIPAC-CM) under Grant Y2018/NMT-4913.

**ABSTRACT** In this paper we present a colour calibration framework for cultural heritage preventive conservation. It includes a custom colour chart based on a regular sRGB division and a multi-dimensional calibration approach based on hyper-surfaces that map the original colours to the calibrated space taking information from all the channels. As shown, our work outperforms past proposals. This solution has been designed to be automatic and match the requirements of our cultural heritage surveillance framework needs, such as non-invasiveness and easiness to deploy. Its performance has been compared to publicly available calibration proposals and charts, and their results judged basing on the context and restrictions of the intended goal, such as illumination. Conclusions on the need for adapting technological solutions to the context of the goal, its cases of use when valid and self-designed methods are drawn. This work is part of the MIPAC-CM project.

**INDEX TERMS** Colour chart, colour calibration, cultural heritage, preventive conservation.

## I. INTRODUCTION

Cultural heritage preservation has been a major concern of the worldly digital and information horizon for the past years, being included in the agendas of governmental organisms [1]. Proposals for achieving a successful cultural conservation involving digital technology, have been launched in big numbers [2], especially considering the wide meaning of the term and the different possibilities of its fulfilment.

Preventive conservation is a way of achieving preservation that focuses on safeguarding heritage integrity, and opens for a multi-disciplinary approach [3]. Preventive maintenance means taking care of the concerned piece before any kind of damage happens. It is also desirable that the prevention system is the least invasive to the piece as possible. This implies that its presence should, by any means, not damage the piece, nor having physical contact with it, and being as less noticeable by third parties as possible. This leaves space for any kind of monitoring deployment involving a

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Bouchir<sup>3</sup>.

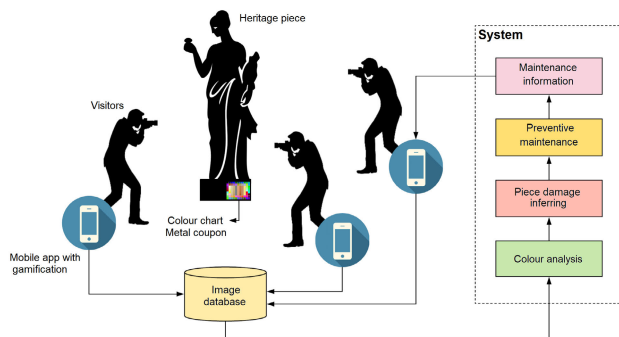
variety of sensing-related approaches [4]. Therefore, when involving technology, preventive maintenance is likely one of the most favouring fields for blending cultural heritage with it [5]. Ambient condition sensors are amongst the most popular ones for cultural heritage control, like humidity, light, emissions or other chemicals that may affect the state of the concerned pieces [6], [7]. In addition to these, image acquisition related means are also proposed, as image data are considered as valid for heritage preservation as a non-invasive method [8].

Colour is an important visual cue for assessing the state of many materials [9]. Metals like iron, copper, silver or lead change their surface colour when exposed to degrading ambient conditions [10], [11]. Thus, visually checking the appearance of any material prone to degradation is a valid sensor-related surveillance mean.

When collecting colour data as indicative cues, camera disparity has to be considered. Due to subtle physical differences in their construction, no camera acquires colour information the same way as any other [12]. When there is no possibility of employing the same camera for performing all the work,

a colour calibration scheme is needed, so the information, regardless of the source, can be homogenised. There exist well-known, reliable and available calibration systems that can be self-programmed or purchased and installed [13], alongside generally employed open-source solutions [14]. Nevertheless, when recurring to these, their conditions of usage may not always be adequate to the followed goals, such as the need of manual handling, or the installation of additional software or physical assets that are out of the scope. Besides, when purchasing a commercial solution, the price could be prohibitive if the dimension of the monitoring system is of a considerable size and requires of a big amount of assets.

In a previously published paper, we introduced a cultural heritage preservation framework based on image crowdsourcing [15] (figure 1). Visitors to the cultural heritage exhibition site are encouraged (and also stimulated through gamification strategies [16]) to take pictures of metal dosimeters and other materials sensitive to degradation placed next to the pieces under surveillance, which will act as indicators of the presence of degrading agents in the site, basing on their surface colour. Considering that all the third-party cameras will be different, a colour calibration is needed so the information of each crowdsourced image is homogenised and reliable. In our previous work we proposed a calibration scheme based on linear interpolation in the sRGB space layers of the acquired images that improved other state-of-art proposals with similar goal, relying on a set of reference colours present on a chart additional to the dosimeters. This is a crucial part of the MIPAC-CM project.



**FIGURE 1.** Schematic view of the project in its global scope.

In this paper, we present the natural continuation of the lastly proposed framework. The linear calibration approach is replaced with a multi-dimensional scheme that maps sRGB values present in an image to the calibrated space accurately employing information from all three dimensions to calculate one channel, effectively forming a hyper-surface as transfer function. Its calculation is unique for every concerned image, and does not require manual control nor the installation of a specialised software package.

Aside, we designed a particular colour calibration chart to be employed in the context of this project, that partitions

the sRGB space in regular intervals to be mathematically generalist. It is conceived to be easily reproducible, cheap and ready to be deployed in big quantities, whilst allowing a possible reduction of its invasiveness towards the piece under surveillance when needed by employing a subset of the colour selection.

It is also an intention of this paper, aside to present a valid improvement to our previous work, to assess that generally available colour chart solutions do not always work so well when having to be adapted to the particular circumstances of certain applications, and that the design and deployment of self-made calibration techniques, specifically designed to deal with the particular problematic in question, are feasible and offer good results.

## II. RELATED WORK

When gathering data, crowdsourcing is an interesting approach due to its scientific possibilities [17], including ways of involving it with cultural heritage preservation [18].

Recent research with crowdsourcing aimed to help preservation has been done, with successful results [6], [19]. Control and monitoring can be achieved by having a constant feed of third-party data about the status of a piece. Third-party data coming from cellphones is also considered. [20]. This also fulfils the non-invasiveness requirement [6].

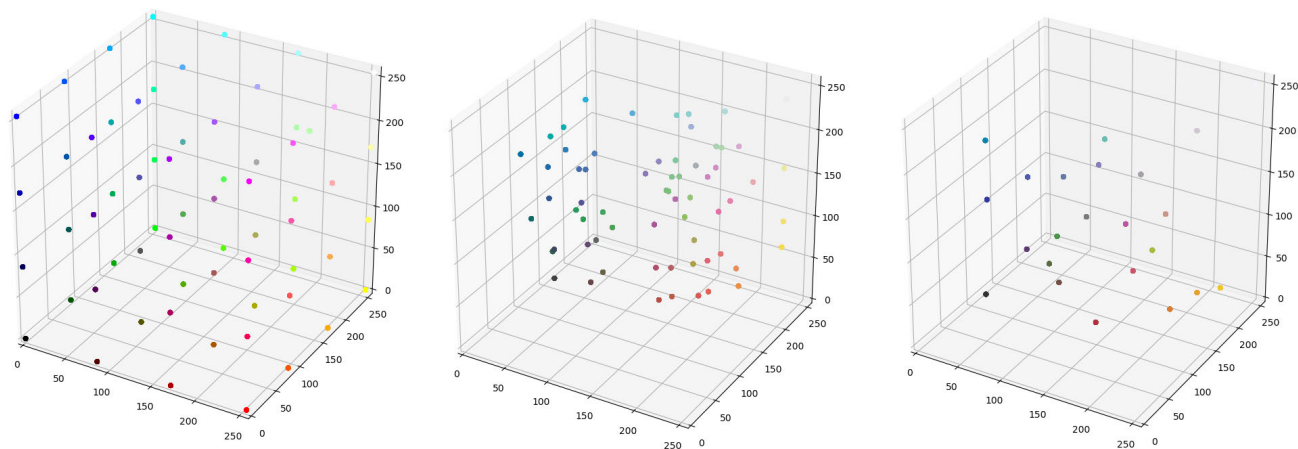
When involving cultural heritage in controlled environments such as museums, lighting conditions are important to assess the visual status of the exposed materials. Whilst there exist some optimal guidelines on how to perform illumination, there exists no standard that imposes a particular criterion [21]. This is a problem that calibration also addresses, given that different illuminants affect on how colours, e.g. in a colour calibration chart, are perceived [22], [23]. Some approaches, especially agriculture-oriented, perform calibration to overcome the differences between sunlight in different times of the day when checking plants [24].

Solutions for colour calibration require the presence of two fundamental factors aside from illumination: the colour calibration chart and the mathematical basis for the proper calibration process [25]. Professional and widely available proposals may include both, offered by companies such as X-Rite [13], [26] or Imatest [27], or a specific software meant for fulfilling the operation, like Fiji [14].

However, some works take advantage of the mathematical basis of publicly available calibration schemes, such as Color Calibration Matrix (CCM) [28], and adapt it to optimise the process to their goals. Proposals to calculate the calibration matrix range from simple processes like linear algebra approaches [24], to other more elaborated ones involving least squares [29] or deep learning [30] have been published. Aside, other works also present custom color charts for the contexts of their own projects [31].

## III. MATERIALS AND METHODS

In this section, we describe the reasoning behind our designed approaches and assets. These should deliver good results



**FIGURE 2.** Views in space of the distributions of different colour charts: Ideal sRGB partition of the self-made chart (left); spectral values of the printed partition (centre); values of the X-Rite commercial chart (right).

whilst maintaining themselves in the appropriate conditions needed to achieve the goal in its pretended context.

**A. COLOUR CHARTS**

In order to perform a correct colour calibration process we need to deploy a colour target that needs to appear in the pictures to be corrected. Usually, this kind of targets offer a set of printed patches whose corresponding RGB values are the “anchors” or references that serve as basis for building the mathematical calibration operation.

Most colour calibration applications in professional and amateur fields rely on commercially available charts whose good performance has been proved to be useful in most cases. One of the most widely recurred charts is X-Rite’s ColorChecker Classic/MacBeth ColorChecker Chart, conceived to represent frequent colours existing in average real scenes [26]. There exist other variants of the ColorChecker Classic, including a wider amount of luminance-oriented colour values and a more exhaustive shading of the basic featured tones, such as ColorChecker Digital SG.

However, even with their conception of being multi-purpose oriented charts, there is always the question if commercially available targets like X-Rite’s can be the most suitable ones to be found for any general or specific application. This question is interesting, especially considering that factors like source illumination and the colour reference selection can be decisive [24]. It is also a factor to consider that its array of featured colours is conceived with a perceptual and not a strictly mathematical criterion.

Thus, we conceived a self-designed colour chart with the goal in mind to display mathematical regularity. Considering that the sRGB colour space takes form of a Cartesian “cubic” three-axed space, we divided each of its axes into regular segments, and the colour combinations resulting of that division are taken to form the chart. The selection of colours is decided upon the number of patches that can be deployed and conform

the set derived from equation (1), and also the acceptable physical size of the chart so it does not result invasive.

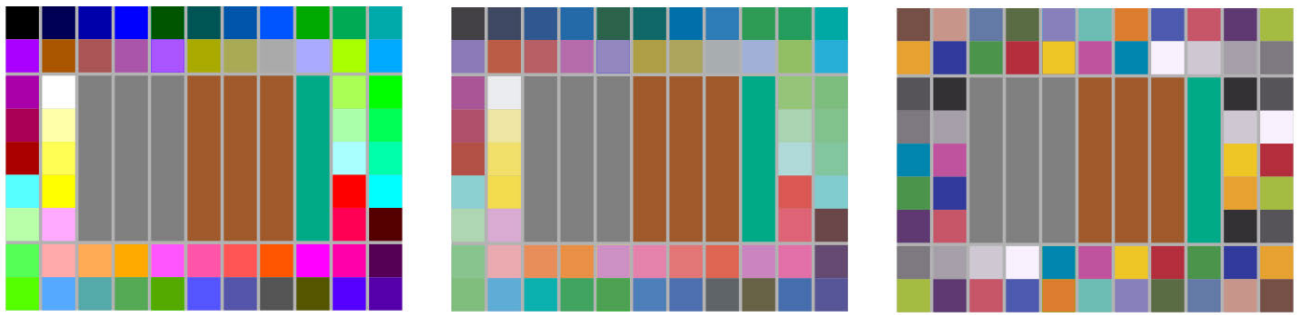
$$S_{sRGB} = \{(R, G, B) \mid 0 \leq R, G, B < 2^{n\_bits}; R, G, B \bmod step = 0\} \quad (1)$$

Here  $step = 2^{n\_bits} / (\sqrt[3]{number\_of\_patches} - 1)$  and  $n\_bits$  is the number of bits per colour used for quantization in the colour space. Note that  $\sqrt[3]{number\_of\_patches}$  should return an integer, so space is covered homogeneously by the division in all three dimensions and all patches fit in space.

If  $number\_of\_patches$  is adjusted to 64, it is possible to generate a physical colour chart as big to host enough large colour patches that will give reliable cues of their sRGB values, but small enough to be placed next to heritage pieces without bothering their inspection. Also, if adjusting and 8 bits per colour,  $(R, G, B) \in \mathbb{Z}^3$ . Lastly, if decimal sRGB values should be returned by the equation,  $(R, G, B)$  values should be quantized to the nearest value if needed.

Effectively, the conceived chart consists of 64 patches that correspond to the colours resulting from the commented division, given the aforementioned advantages and the feasible dimension requirements. Our design also includes the “corners” and limits of the sRGB space, so all the RGB and CYM primaries are included, as well as black and white tones (figure 2).

There is another consideration to be taken for the chart design. When performing a colour calibration, it is expectable that the colour reference values can be found in real environments when the pictures to be calibrated depict real scenes. This is especially important when concerning the extremes of the domain, like the primaries or pure black and pure white. Normally, real scenes will feature a limited gamut, spanning the interior of the total mathematical range of perceivable colours in space and leaving the extremes aside. Therefore, it is important to assure that the featured hues in the designed chart depict sensibly real colours, too.



**FIGURE 3.** From left to right: (1) initial sRGB value division; (2) self-designed colour chart based on the spectral values of (1); (3) colour chart based on X-Rite's ColorChecker Classic.

In order to achieve this, the “real” corresponding values of the ideally conceived division have to be found, and then the chart needs to be built with them, so a sensible and accurate calibration can be achieved (figure 3). So, the values of the aforementioned ideal sRGB space division have been printed with the same materials and conditions that were chosen in reference [16] for the physical chart to be deployed in real environments, and only then their spectral values measured with a Konica Minolta CM-700-d spectrophotometer. These measurements were then transformed to the sRGB (figure 2). Therefore, now the calibration chart with the appropriate reference values can be printed and deployed. Despite the fact that the printing process will inevitably reduce the fraction of sRGB space covered by the colour chart, having the sRGB range fully covered in its conception means that the range reduction of the real-life colours will be less exaggerated if only the interior fraction of sRGB was considered in the initial steps of the design.

The self-designed colour chart has been printed upon acid-free laminated paper, using an industrial ink set LED-UV Xtreme Pro, as stated in the previous article [16]. In this work, we described why these materials are suitable for producing colour calibration charts for the aim of this project, namely cultural heritage monitoring in a controlled environment. This also allows for a quick, easy and cheaper chart reproduction for many simultaneous usages and calibrations, so a big amount of them can be deployed or substituted at the same time. It has an advantage compared to other commercial charts in this sense, where their price hinders their usage in a massive way such as monitoring a big amount of pieces. These are thus being more adequate for a reduced number of non-simultaneous calibrations.

In the following, however, the calibration performance for both colour selections, X-Rite's commercial and our own chosen mathematical criterion, are tested. In order to regularise the charts, we have printed the colour hues of the ColorChecker Classic the same way we printed our self-designed chart, arranging them in an analogous way. Given that XRite's chart features a reduced number of colours compared to our design, we added patches with repetitions of the same colours, so the same visual design of the chart can

be achieved. When performing the calibration, all of them or only a subselection can be selected. Repeating reference values included in the calibration processes has shown having no influence, furthermore.

The colour patches in the charts are arranged on a pattern of 9 rows and 11 columns. In the central part of them, a space of  $5 \times 7$  patches has been reserved for the aforementioned dosimeter material tags. These will be the monitoring cues for performing an effective surveillance.

## B. CALIBRATION TECHNIQUES

In our previous article [15], we presented a colour calibration approach specifically designed for the followed context and goal, based on linear interpolation for the range  $[0, 255]$  of the sRGB R, G and B layers separately. Given the assumed discordance between cameras of origin, it has been conceived as an adaptive technique bound to overcome it and represent the image information in an homogeneous domain, regardless of the source. Thus, it calculates its parameters from and for every single supplied image exclusively.

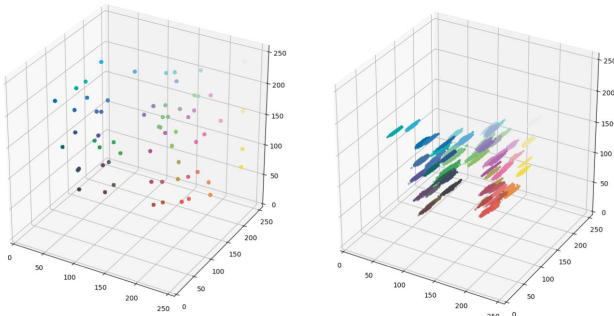
In this paper we present an improvement to the previous linear calibration. Following the same basic idea, the concept is expanded so the sRGB information for every pixel in the R, G and B layers is taken into account to calculate each of the values for every coordinate in the calibrated space. This is equivalent to calculating three hyper-surfaces that relate three-dimensional information to one dimension each time separately. Analogue to the previously proposed linear calibration, each hyper-surface acts as a transfer function that maps all the sRGB domain to the calibrated RGB values. The resulting one-dimensional values of the three transfer functions are then stacked, so the new R, G and B layers of the calibrated image are formed.

This multi-dimensional calibration divides itself in two fundamental sequential steps. Firstly, the colour reference data are calculated with the values from the chart using a variant of the Reinhard color transfer technique [32], so the position of the reference data in the sRGB cube and their relationship with their homologues in the pictures to be processed can be described. Secondly, the construction of the R, G and B spatial transfer functions is performed, so the value mapping



can be performed. Both steps are entirely calculated using the single image to be processed, knowing the set of reference colours in beforehand.

If we represented the colour patches out of the chart in the sRGB space, ideally they should take the form of regularly scattered dots along the Cartesian “cube”. However, when taking any image representing the chart to be processed, each patch will feature some visual heterogeneity among its pixel values due to the effects of the inks when printed, the angle of light reflection on its surface, the digital camera sensor response, etc. What was originally conceived as an uniform, single colour value printed in a small physical region, is perceived as an approximately Normal distribution surrounding an average value when digitalised (see figure 4). Therefore, the shape of the colour patches in a digital image, when represented in the sRGB space, will be seen as a group of three-dimensional point clouds that also have lost the regular distribution intended with the initial sRGB division. Shortly said, the colours of the chart are displaced and instead of being described by an unique position  $(R_{ori}, G_{ori}, B_{ori})$  in the R, G and B axis, they now possess a distribution of  $N(\mu, \sigma)$  characteristics in every coordinate,  $(N_R, N_G, N_B) \neq (R_{ori}, G_{ori}, B_{ori})$ .



**FIGURE 4.** Views in space of the distributions of the conceived colour chart of reference (left) and the effects of printing and taking a picture in real conditions over it (right).

Taking this phenomenon into account, the first step of the multi-dimensional calibration consists in recovering the spatial statistic characteristics of the ideal sRGB distribution, calculated from the colour patches in each image to be processed. This will enable the possibility to build the “skeleton” of the transfer functions as anchor points of the hyper-surfaces to be constructed. In order to achieve this, a modified version of the Reinhard color transfer technique is applied. The Reinhardt color transfer consists in modifying the mean and the standard deviation of a numerical distribution, so it is displaced in space and its width is modified, but without actually losing information on peaks and maintaining proportions in space. When applied to the colour histogram of the R, G and B layers of any given image, it effectively produces a colour shift of the overall tone in it.

The effect of the Reinhard transfer can be used for producing an equivalent mathematical distribution shifting in 3D.

Instead of modifying an image layer histogram, each one of the digital colour patch distributions of the chart in space can be shifted in three dimensions (equation (2)), so the positions in space of the ideally conceived chart can be regained parting from its appearance in the images to be processed.

$$\begin{pmatrix} R_{out} \\ G_{out} \\ B_{out} \end{pmatrix} = \begin{pmatrix} R_{in} - \mu_{Rin} \\ G_{in} - \mu_{Gin} \\ B_{in} - \mu_{Bin} \end{pmatrix} \begin{pmatrix} \theta_R & 0 & 0 \\ 0 & \theta_G & 0 \\ 0 & 0 & \theta_B \end{pmatrix} + \begin{pmatrix} \mu_{Rref} \\ \mu_{Gref} \\ \mu_{Bref} \end{pmatrix} \quad (2)$$

Effectively, this matrix operation means a displacement and reshaping of every distribution in space. If the target colour distribution consists of a single point ( $\sigma = 0$ ), the standard deviation is left untouched, so the value range of every patch distribution is respected and there is no loss of image pixel information in this sense (equation (3)).

$$\theta_i = \begin{cases} \frac{\sigma_{iref}}{\sigma_{iin}} & \sigma_{iref} \neq 0 \\ 1 & \sigma_{iref} = 0 \end{cases} \quad (3)$$

When finishing this first step, every pixel value corresponding to the colour patches in the chart will have a corresponding homologue rearranged in the sRGB space. These will conform the anchor points of the multi-dimensional transfer functions to be built, and thus the “skeleton” of the calibrated sRGB space to be calculated. The following process consists in finding out the rest of values between the anchor points, by connecting them in all three dimensions.

To achieve this, the hyper-surfaces  $f(x_1, x_2, x_3)$  are calculated analogously to the firstly proposed linear calibration (4), but extended to the amount of dimensions needed for this application ( $N_{dim} = 3$ ). This process is based upon a trilinear interpolation (5), where  $P(r_k, g_k, b_k)$ ,  $k = 1, 2$  stand for the anchor points of the cube that surrounds the sRGB space region where the hyper-surface is constructed. These anchor points are composed by combination of two different r, g and b coordinates.

$$y_p = y_1 + (x_p - x_1) \frac{y_2 - y_1}{x_2 - x_1} \quad (4)$$

$$f(x_1, x_2, x_3) = \sum_{k=1,2} P(r_k, g_k, b_k) \prod_{i=1}^3 \begin{cases} x_i & k = 2 \\ (1 - x_i) & k = 1 \end{cases} \quad (5)$$

Lastly, having built the three transfer functions following this criterion, the R, G and B layers of the calibrated image are calculated by projecting every pixel of the images to be processed using them.

In the following sections, the performance of the multi-dimensional calibration is compared with the previous linear scheme and its stronger points objectively evaluated. Aside, in order to achieve a more complete comparative experiment, the results obtained with a widely utilised scheme such as Colour Correction Matrix (CCM) are also featured. Normally, the CCM calibration is favoured to be performed with the help of specialised software tools, which

allows a manual parameter selection for an elaborated optimization process [28].

Nevertheless, for the sought kind of cultural heritage monitoring, this method is infeasible. Surveillance of the pieces in question should be automated and adaptive, considering that the amount of crowdsourced pictures can reach big numbers and involving a permanent human effort would mean additional undesired costs. Having this restriction, the CCM calibration to be tested here needs to be adapted to the conditions of this project. Other works also do their personal approaches to the basic CCM idea [30]. Thus, a resolution of the CCM based in linear algebra, following the reasoning in [24], would be more feasible, quick and appropriate for the comparison to be explained in this paper.

The process for calculating the CCM here is thus unique for every image. First of all, representative values for the colour patch distributions for each image to be processed are extracted. These are the means of the distributions. Parallel to this, the ideal values of the colour reference patches are also extracted. Using both sets of values, a  $3 \times 3$  matrix is deduced (6). The calibration is achieved by multiplying every image pixel with it.

$$R = CCM \times O_r \quad (6)$$

In equation (6), it is stated that the set of reference values  $R$  should be equivalent to the CCM multiplied by the set of same colours in the original image  $O_r$ . Elaborating on this, CCM can be calculated by equation (7).

$$CCM = (R \times O_r^T) \times (O_r \times O_r^T)^{-1} \quad (7)$$

This method also ensures saving colour distribution information by assigning a different calibrated value to every original value. However, it should be performed in the mathematical linear RGB space, so the sRGB associated gamma non-linearity does not affect the results of an otherwise linear operation. Thus, the CCM calibration approach requires a gamma preprocessing and postprocessing so the sRGB information stays coherent with the calibration and the linear operations maintain their linearity.

#### IV. EXPERIMENTAL

The chart and the mathematical calibration basis, as the two critical parts of the scheme, have therefore been tested in order to measure their performance quality in the context of this work.

The case of usage consists in collecting crowdsourced pictures from different cameras of origin. Our last proposal addressing this topic [15] showed the robustness of the linear scheme when dealing with the inter-camera acquisition variability. Therefore, it is also adequate to test the other calibration schemes under different image conditions that might alter the colour and visual characteristics.

Henceforth, four sets of test images have been prepared to check the performance of the linear, multi-dimensional and CCM calibration schemes employing the previously prepared two variants of the colour charts. The dosimeter placeholder

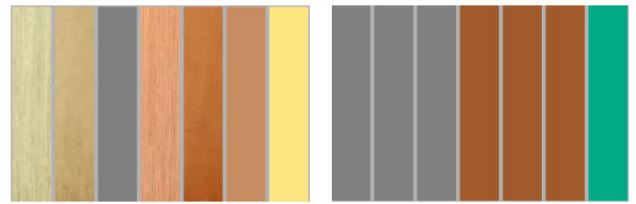


FIGURE 5. The two types of test dosimeters employed.

space, in the middle of the targets, has been substituted with uniform colour stripes that imitate the average tone of copper, silver and PH-sensitive paper, and different textures that have been taken of real photographs of copper and silver tags (figure 5). So, the effect of the calibration and accuracy can be measured on plain colours different from the anchor points, on the gray luminance-related region of sRGB space, and on approximately realist colour information of the real targets to be processed in field cases. Therefore, four different kinds of test charts have been printed.

It is utmost important to consider the effect of the calibration on the dosimeters. Their colour values, and especially their colorimetric differences with the original, non-degraded state, is essential for performing an effective surveillance on the state of the cultural heritage pieces. Then, the calibration performance on their equivalent textures has to be evaluated in a similar way.

This is the reason why the success of the tests has been measured with objective means. Given that the monitoring of the colour change in the dosimeters is meant to be done in the measurement-oriented CIELAB space, the calibration performance on their placeholders is evaluated by taking the average CIEDE2000 [33] difference of the calibrated picture compared with the equivalent ideal dosimeters in the digitally designed colour charts. This is the main metric taken for judging quality, and the one used as a guide for the evaluation. Apart, CIEDE94 [34] has been taken, to see its effects on an alternate metric. The results of these can be found in the additional tables (Results set 1, Results set 2, Results set 3 and Results set 4) supplied with this article.

Lastly, and as a mean to assess how accurate each calibration approach is, the same distance measures are taken on the colour chart reference patch distributions. This way, the closeness of the transfer functions with its initial “skeleton” made with the original anchor points can be evaluated.

The four image sets feature acquisition of the charts under different lighting conditions, ordered from the lowest to the highest illuminant temperature. Set 1 consists in pictures taken under a common tungsten light bulb, with a yellow-orange light of low brightness. Set 2 was taken under direct sunlight at noon (white light with elevated brightness), whilst sets 3 and 4 were taken under common fluorescent illuminants of high temperature (blue light), of high and low brightness each. The particular values of illuminant temperature are left unknown, as the intention of the project is to

TABLE 1. Variable factors of the calibration experiments performed for each image set.

Calibration	CCM		Multi-dimensional		Linear		
Chart	X-Rite			Self-made sRGB division			
Dosimeters	Uniform			Textured			
References	All patches	RGB + CYM + B + W	RGB + CYM + W	RGB + CYM	RGB + B + W	RGB + W	RGB

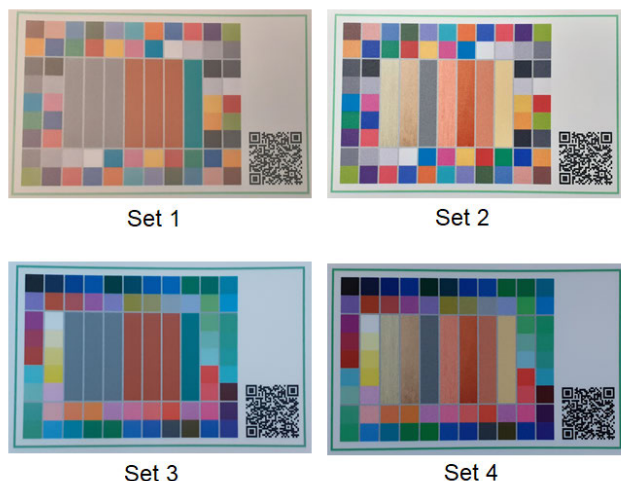


FIGURE 6. Example images of the four sets considered. The different lighting temperature and brightness is to be noted.

test the performance under any condition, given the lack of standardization on interior lighting [21]. All of them were taken with a conventional cellphone camera that codifies the information in sRGB, as the vast majority of crowdsourced pictures from conventional cellphone cameras are expected to be [15].

According to this, it is reasonable to think that the physical placements of the charts under monitoring might incur in slight illumination differences depending on the room or the exhibition, which is possible to happen even in the same room depending on their location. Thus, the selection of test images covers both the effects of inter-camera acquisition variability, with different perceptions of the same colours, and the visual aspect provoked by non-regulated lighting sources.

Aside, it is very interesting to check the calibration performance when using a reduced subset of the original anchor colours for building the transfer functions. Some papers [24] state possible benefits and better performance depending on the number of reference values to take on calibration depending on the context of usage, namely on reducing it. This is also a fundamental question to address when considering that the colour chart for monitoring cultural heritage is an invasive element among the exposed pieces, so any justified reason for its presence reduction would be adequate to check, too. The reduced amount of material implies an even cheaper cost of chart reproduction and deployment.

Henceforth, the tests have been performed for the complete set of colour patches, and for reduced subsets consisting in the CYM and/or RGB primaries (or their closest equivalent),

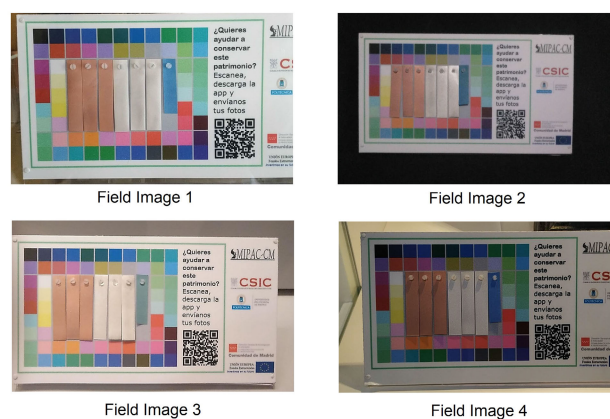


FIGURE 7. The four field images employed to assess the performance of the chart and calibration on real cases.

adding black and/or white. This has been repeated for all images in the four sets of different characteristics, and in both kinds of produced charts with two different dosimeter texture testings. Table 1 shows the variable factors that have been combined for experimenting with each set of images.

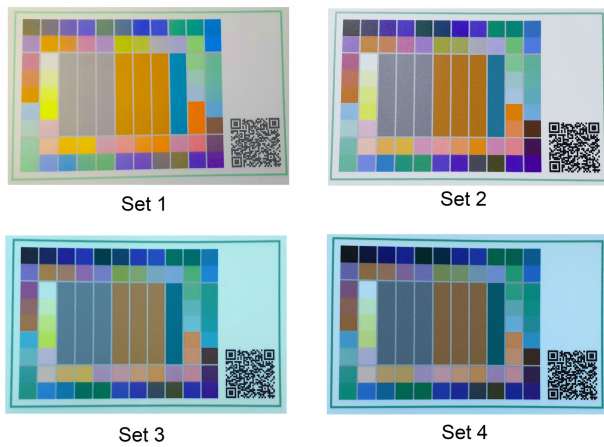
Lastly, so the effectiveness of our scheme can be assessed to the last extremes, the explained experiments have also been executed over field pictures aside from the test sets. The self-designed chart has been printed in the aforementioned procedure and deployed in a museum during an one-year test period, and therefore some pictures taken as a testing ground. Four representative pictures of different chart placements have been chosen, shot from different distances and showing different, uncontrolled illuminations (figure 7). Therefore, the sample size has been of 16 test images (two anchor colour selections and two kinds of dosimeters in four sets of images with different illumination), and 4 field images.

## V. RESULTS AND DISCUSSION

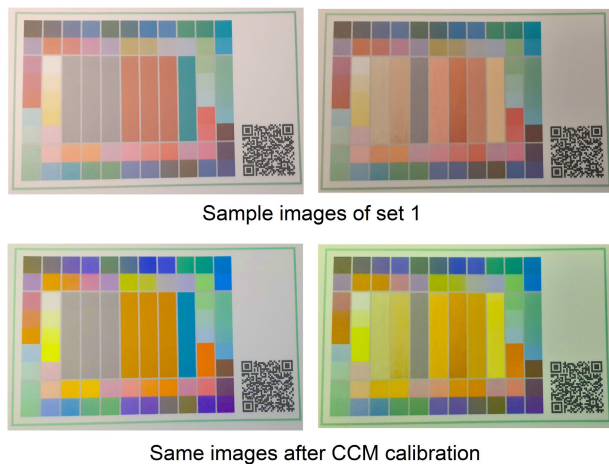
The tests have been performed on the pictures of the four considered sets and field images. The visual appearance of the non-calibrated pictures corresponds to what is expected from real crowdsourced information by conventional cellphone cameras: different illuminations and shadings, pixeling effects in what was initially conceived as uniform colour regions and overall different colour tones.

This characteristic in the pictures has been observed as being considerably influential over the calibration process. This is an especially prominent effect to be observed on the CCM calibrations. After going through the process, the overall colour tone given by the illuminant is exaggerated





**FIGURE 8.** Example images of the four sets considered after the CCM calibration.



**FIGURE 9.** Variable effects of the CCM calibration with images from the same set.

(figure 8). Thus, all pictures from the Set 1 after going through the CCM present an exaggerated yellow-green tone shifting, whilst sets 3 and 4 are shifted towards blue-violet. Set 2, given its white illuminant, does not feature that effect. When observing each picture in detail, it can be seen that the overall tone shifting is slightly different for each image of the same set, which is indicative of the calibration being unique for every case and not a generalised effect (figure 9). It is understandable to think that, given that the CCM calculation is made by means of taking the average value for each anchor colour distribution, and due to its pixeling effect, it will be slightly different in each picture, even among the same set under the same illuminant.

The colour tone shifting does not happen with the self-conceived linear and multi-dimensional calibrations. However, these incur in another different effect that is not to be found in the CCM results. Whilst CCM, being fundamentally a multiplication with a certain value, respects the pixel

distribution variability, the other two techniques may lose that coherence in some cases. This is especially predominant in the linear approach, where every layer transfer function and therefore coordinate is calculated separately, and thus the resulting pixel coordinate values might have not the same proportion than in the original image, effectively assigning “false” colours to them. The multi-dimensional approach minimizes this effect by extracting information from all layers for every transformation. Curiously, this effect is more prone to happen in tests where a bigger amount of reference colours is taken. This happens due to the different values of a pixel distribution falling into different interpolation regions that are close in space, so the bigger the interpolation region, the less distribution alteration happens.

Nevertheless, given that the final intention is to measure the calibration quality by objective means, the objective distance metrics  $\Delta E$  mentioned in the Experimental section are taken as representative to the success of the designed scheme. This is also the most reliable way to assure that the calculated transform functions are accurate, and to compare the performance between charts.

First of all, the accuracy of the calibration schemes needs to be measured. In order to do that, mean CIEDE2000  $\Delta E$  colour differences between each colour patch of the reference data and calibrated images are evaluated. Table 2 shows the mean and the standard deviation for the difference metrics for all the 64 colour patches taking the calibrated images from all the four sets, for every calibration approach. In all cases, the multi-dimensional calibration offers a mean metric between 0 and 1, indicating an extremely accurate resemblance, almost unnoticeable, with a very slight variation. This indicates an accurate structuring of the “skeleton” of the transfer functions. These metrics are superior to the CCM and linear approaches, where the effect of the colour shifting and the false pixel colour assignments respectively can be observed by the big differences and the stronger variability.

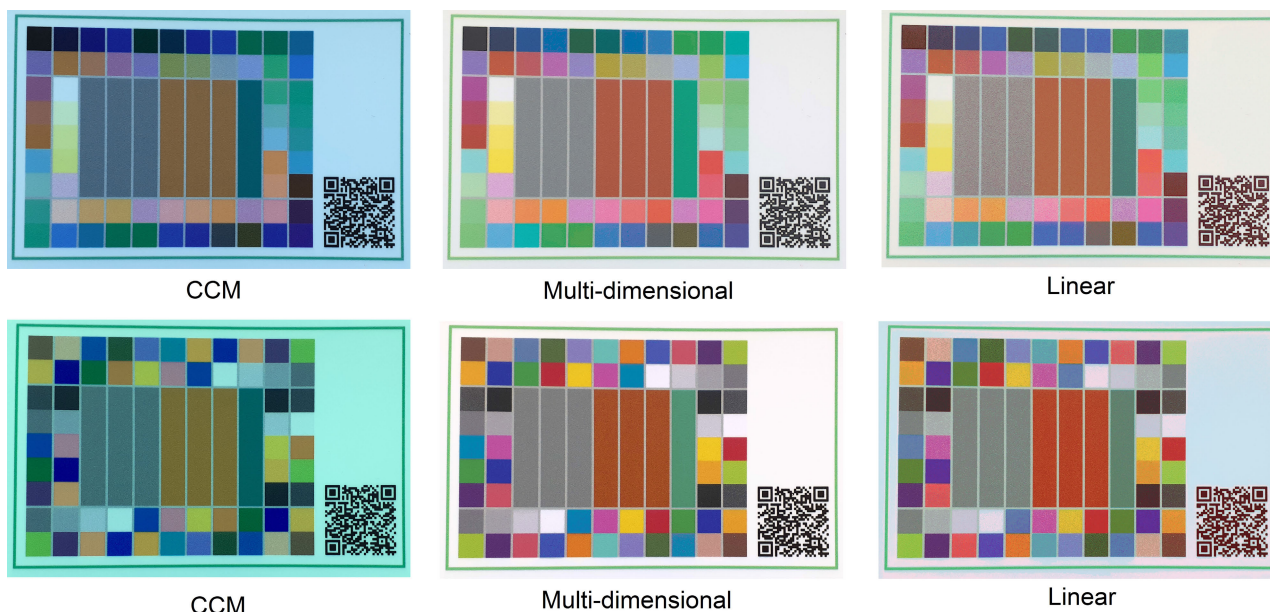
Then, the mean difference values between the calibrated images and the reference data are calculated for the dosimeter textures. Here, a comparative study for all the images inter and intra-set has been performed, observing the performance for the self-designed and the X-Rite-inspired colour charts. Two fundamental criteria have been taken: accuracy and its improvement between both kinds of charts, and robustness against reduction of the anchor colours.

Thinking logically, it may appear that the more amount of anchor colour values taken to build the calibration scheme, the better the performance is meant to be. This conclusion is obvious to reach when considering that more points means a bigger range and a finer division of the space, a fact that should imply a more accurate value mapping when building the curves and a lower probability of getting an exaggerated interpolation error. Nevertheless, some studies [24] show that not only reducing the amount of anchor colours to a specific subset does not only not affect the calibration quality, but it also may improve it when using CCM and X-Rite’s colour chart. This kind of result is not as sure to happen when



**TABLE 2.** Accuracy CIEDE2000  $\Delta E$  metrics of the 64 colour patches of the self-made colour chart for the three colour calibration schemes. It can be seen that the multi-dimensional scheme, in columns 2 and 9, offers the best similarity.

Patch	Multi-dimensional		Linear		CCM		Patch	Multi-dimensional		Linear		CCM	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.		Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
0	0.412837	0.151512	14.08954945	3.059626732	16.76511253	6.358449886	32	0.194245	0.068097	5.099936446	2.027045382	14.38076683	7.395351129
1	0.260016	0.095811	9.216607357	3.594088393	18.69849987	3.474201697	33	0.1956	0.080486	3.441437826	1.652833378	14.63217168	4.037755691
2	0.228134	0.079358	8.108327153	4.505489788	15.61821867	4.07554405	34	0.223439	0.05488	7.059545631	2.878427889	16.77922446	5.497386578
3	0.310319	0.125786	9.914985233	4.528780565	27.20284731	5.29005594	35	0.206441	0.162759	9.473714974	5.173247062	10.28514833	3.369350842
4	0.170674	0.072103	5.026345089	3.101142757	13.49644562	6.645231737	36	0.217311	0.092866	6.647084603	3.238750017	16.39197341	6.6210191
5	0.197266	0.080123	8.935710313	3.660609057	27.23512208	5.064481151	37	0.199861	0.076772	6.565954133	2.376879838	15.2785667	5.785934768
6	0.204217	0.03479	4.93815658	1.740167236	15.04732521	5.901447917	38	0.210568	0.068017	4.936360059	4.71962697	11.77188431	4.037855603
7	0.173662	0.068409	7.407301417	2.363916263	17.79188931	4.584355027	39	0.374855	0.425229	9.049664351	5.172686372	9.30930693	3.617544662
8	0.309231	0.057458	6.85689961	5.474068195	13.77471987	2.228122922	40	0.251445	0.088826	4.93619238	1.384245492	16.53917243	4.735062478
9	0.292487	0.062035	10.37152028	3.533123239	13.92500883	2.238462358	41	0.234356	0.106102	5.069210301	2.079133739	14.72922234	4.998865804
10	0.268096	0.095789	7.053852694	2.220406788	12.90858587	3.708256286	42	0.30284	0.113546	4.777480353	2.447513911	11.95111392	4.266366286
11	0.582764	0.822917	5.533874808	1.419840192	20.19448379	3.559907877	43	0.256648	0.117357	5.557981745	3.163952444	9.756538947	3.875667137
12	0.183796	0.057675	3.703390895	1.978333338	14.57789976	7.297762122	44	0.203413	0.06507	4.416594532	2.146541352	13.11720793	6.063364684
13	0.210032	0.087019	7.576334888	1.858891576	25.05028155	5.276838651	45	0.205403	0.059684	4.345215922	2.134475368	12.46753773	6.74331365
14	0.180852	0.063165	4.058300036	1.587937988	14.01199788	7.587729878	46	0.145512	0.06191	2.693377103	1.71619415	13.40205788	5.53927612
15	0.155289	0.070665	4.249909504	2.756259784	13.61027617	6.199758287	47	0.224507	0.097113	3.638742323	2.745851102	13.26553805	3.651625766
16	0.184675	0.091652	5.402757314	2.64497672	15.45579863	6.688914326	48	0.17285	0.067303	4.190340846	1.130007527	15.83154039	2.258199194
17	0.188428	0.064692	4.665882909	3.152100089	13.19538794	4.49264332	49	0.224538	0.08082	3.993770567	0.742405989	15.68386475	2.67982209
18	0.37635	0.178005	13.23543682	4.990741384	14.37898698	5.866363369	50	0.198844	0.081617	3.450194147	2.002333558	13.22215485	2.689562084
19	1.544123	1.328671	11.08715949	4.733126074	23.15314444	3.773997003	51	0.220565	0.079902	5.653809655	2.808151186	12.16084103	4.459272473
20	0.198739	0.112302	8.827212102	3.724085625	20.47258809	5.068975051	52	0.210006	0.074186	2.972181562	1.496902769	14.46974157	2.83536853
21	0.230035	0.097683	9.855765879	3.41298248	23.654733	5.362597386	53	0.204534	0.075385	3.099102181	1.275749874	13.34710559	3.471615339
22	0.169928	0.058807	4.919968286	2.536362927	15.61931619	6.223550472	54	0.208231	0.050929	5.36897584	1.796432616	13.96406747	2.95146949
23	0.179962	0.074765	6.429902286	3.864490041	22.85988777	5.911263806	55	0.19878	0.059795	7.597487754	3.546060564	12.82476095	4.793131873
24	0.180385	0.047244	4.991662837	1.300609839	15.8224371	5.421694186	56	0.131247	0.047172	4.394161581	2.118888528	17.28307394	2.06904037
25	0.222874	0.073458	3.544663922	1.682422308	16.79224223	3.754517464	57	0.181062	0.068652	3.886493186	2.106296612	15.15360512	2.947264461
26	0.13065	0.083973	8.332942723	3.083183804	11.21297903	3.623737554	58	0.165495	0.040394	5.1093628	2.092092273	17.61789851	5.158311363
27	0.487766	0.693288	4.328365115	1.061726975	21.73485899	3.798826127	59	0.199418	0.04089	6.18092928	1.569804417	17.16986061	5.638518973
28	0.504643	0.162159	10.71950192	1.97664423	12.76391718	3.606427849	60	0.131282	0.065707	2.10944634	0.7329191	18.94135344	5.225843375
29	1.667529	1.725776	6.811215415	2.903130993	22.4946089	3.582081347	61	0.130962	0.074673	2.821960183	1.581040933	18.16423881	5.518042278
30	0.230211	0.031835	4.430975138	2.370853733	15.98721555	7.178988059	62	0.207867	0.103698	3.93463388	1.192246592	16.07062773	6.539204301
31	0.195985	0.07469	5.601555947	2.248312295	18.27849339	3.822514982	63	0.198705	0.188878	5.108091349	1.98735289	15.2589886	5.32678106



**FIGURE 10.** Comparative view on the effects of the three calibration schemes on sample images of set 4.

employing the self-designed calibration schemes, considering their different mathematical basis.

The effects of the reduction of the amount of anchor colours is measured with the standard deviation of the differences in the dosimeter textures for all the experiments with different amounts of reference colours. The closer to 0, the less effect the reduction has. This is especially important when dealing with the linear and multi-dimensional calibration schemes, considering the bigger span of the interpolated regions.

Table 3 reflects this study on variability on sets 1, 2, 3 and 4, respectively, in different rows. The standard deviation of the CIEDE2000  $\Delta E$  calibration accuracy values for every dosimeter in each reference variation experiment is shown. It can be seen that primarily and in all sets, for the multi-dimensional calibration scheme the stability of the results is better. Furthermore, our self-made colour chart shows a generally even better stability than the X-Rite reference, with a standard deviation between 0 and 2 in most cases, especially in sets 3, 4 and 1.

**TABLE 3. Standard deviation of the CIEDE2000 metrics extracted for each dosimeter in all the experiments with different amounts of reference colours for all the image sets. It is to observe that the most robust to reduction and stable set corresponds to the self-made chart under a multi-dimensional calibration scheme.**

		Image set 1											
		CCM				Multi-dimensional				Linear			
		X-Rite		Self-made		X-Rite		Self-made		X-Rite		Self-made	
Dosimeter		Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured
1		6.96897464	1.242747427	9.144524263	3.787889388	2.238687182	7.324778877	1.403017415	1.958422883	2.399532194	3.97381021	2.713724416	6.241979495
2		6.88083947	1.362539641	9.645008785	5.315729705	2.402302224	7.95105271	1.605866522	1.469397899	2.965780591	3.421298864	2.908294814	4.618928327
3		6.979517294	2.477500858	10.1033193	8.502357236	2.842531651	3.890272263	1.953685855	2.171632198	3.7662232	2.80476949	2.638759576	3.804650203
4		5.806758755	1.928879448	11.05785153	6.358879201	2.340854434	6.37974353	2.002208683	1.015270172	3.6025979	1.710011436	2.23837322	2.982155517
5		5.779964822	2.52517359	11.77135158	11.28435835	2.51637642	4.707701316	1.610418138	3.262678782	2.88752976	2.28752976	2.852965019	2.852965019
6		5.627597933	1.961089932	12.44201706	8.524341088	3.013192466	4.275873257	1.000550228	1.715983373	2.769766166	4.100068632	4.194352918	3.327837639
7		8.220055005	0.78726155	4.544343567	1.818685036	9.977838688	10.19212848	6.656051018	4.969695332	4.572702682	3.681600913	1.894902301	6.188855686
		Image set 2											
		X-Rite		Self-made		X-Rite		Self-made		X-Rite		Self-made	
Dosimeter		Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured
1		6.772497876	2.41824242	6.558020494	5.837182921	2.258598587	3.34964983	1.832504774	2.404338581	2.84447063	5.440545568	0.603143609	4.886842957
2		6.763406071	3.482897547	6.463756959	5.631491849	2.650876741	3.920136846	1.847373036	2.110388907	2.85473442	2.579342321	0.520271877	3.254169818
3		6.78904243	5.307657907	6.299782603	7.648932346	3.035146391	2.443309734	1.602171666	3.85846382	2.604043206	5.142240082	0.64951428	2.670351976
4		4.565628565	5.793678285	6.353389599	7.169722395	2.240495384	3.729768152	0.493376643	1.422147596	2.214579319	0.908993558	2.172155192	1.541348178
5		4.552827651	4.03587331	6.238638326	6.441258693	2.231815011	3.425445118	0.3871017	0.787903976	2.283614818	2.572911512	2.202994716	2.123517081
6		4.544098346	5.264557483	6.189111668	6.473212059	2.156037532	3.095466935	0.389683653	1.910362589	2.431203037	2.271151904	2.246793075	2.107950018
7		10.34368534	0.67292905	5.081002808	2.741950446	2.362130865	3.857329586	6.003688274	2.500385095	3.675667413	1.882288236	2.365859998	4.346987197
		Image set 3											
		X-Rite		Self-made		X-Rite		Self-made		X-Rite		Self-made	
Dosimeter		Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured
1		3.220604357	2.262196153	0.999349194	0.507824147	2.630449588	5.005363894	1.05365712	3.164797371	4.143188976	2.636519962	1.683575056	2.279445466
2		3.187593258	2.464627137	0.922565351	1.349584022	2.611490106	6.238356833	0.895411792	3.098117852	3.919809588	4.218812258	1.298169259	1.018608225
3		3.252078513	2.74824598	0.996114993	0.942868031	2.790892524	1.042526537	0.982788863	0.676987523	4.353063347	4.91357829	1.088816786	0.95876086
4		4.36495305	3.595467349	6.249471187	3.993590823	2.951114004	5.825172471	0.485477916	0.889979563	0.674267599	3.45840478	1.808247665	1.905345823
5		4.357408328	3.416810622	6.360405283	5.061728203	2.96264732	4.362028092	0.447435275	0.60441535	0.617763342	4.217984016	1.673710591	1.393348994
6		4.399363142	3.542247173	6.384856712	3.917036201	3.070049981	5.738240619	0.400603157	2.061351707	1.371620638	1.525554029	1.592366321	0.841551549
7		5.199804119	0.412993973	1.476065444	0.62368847	2.197811768	6.6697943	5.518330093	1.459474443	9.227253866	2.159410862	1.812374868	2.695037146
		Image set 4											
		X-Rite		Self-made		X-Rite		Self-made		X-Rite		Self-made	
Dosimeter		Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured	Uniform	Textured
1		0.801512981	0.575456821	1.389273659	1.37463872	2.254669145	3.452294787	1.534595077	1.444634379	4.636915613	3.137730041	4.510965033	4.397863263
2		0.78044983	0.801888719	1.376329662	0.93155906	2.236823144	6.016532801	1.284236821	1.262290998	4.474088327	3.961937158	4.517962305	5.26297272
3		0.8471218	0.739110461	1.389132372	1.656903051	2.33868744	2.704672135	1.355637765	1.306987295	4.189174215	3.791823952	4.541493288	5.242348901
4		2.703513874	2.393593089	4.527545257	3.600323757	4.362143759	6.828966108	1.751098438	1.293063978	1.815938001	3.822961221	3.011478263	3.518191311
5		2.711665265	3.202528667	4.532518655	4.998402459	4.30909809	6.277158576	0.748483374	1.21962392	1.790908889	4.781825875	3.021247667	2.070622018
6		2.732497554	2.319893241	4.540789183	3.170093261	4.266005466	6.726611925	4.266005466	0.955997812	1.773154189	5.339432587	2.991328205	3.111809697
7		2.310374976	0.552797904	2.364099422	1.221881273	1.817516453	7.653674433	2.157638312	2.541874411	7.68167757	2.816254219	4.609214076	3.054920507

The most variable results are thrown by the CCM calibration, which is understandable considering the colour shift effect on the calibrated pictures. The linear calibration offers results of mixed and variable quality, which arguably come from the pixeling effect it is prone to provoke. Whilst the illumination difference seems to affect these experiments, its effect are seen to be minimised on the multi-dimensional calibration, where the resulting numbers stay in a similar range for all four sets.

This implies that the multi-dimensional calibration framework is the least affected in results by the lighting conditions and reduction of the reference colour set.

In the supplied tables as additional data (Results set 1, Results set 2, Results set 3 and Results set 4), one for each set of images, the average  $\Delta E$  metrics measurements are given for each single target dosimeter in each experiment with different calibration techniques, as seen in the different tags of the tables (CCM, Multi-dimensional and Linear), using both charts as references. The two column groups on the right state the difference in percentage that the usage of the self-made chart means compared to X-Rite's selection. A negative value means a worse performance, whilst a positive indicates a better one.

When inspecting the CIEDE2000 data results for set 1 (low temperature and brightness), it can be seen that, generally, performance when employing X-Rite's chart is better, in some cases with a vast difference, whilst in the multi-dimensional calibration the self-made does a better result with reduced reference sets, corresponding to the

observed before. The range of differences stays mostly between 2 and 14, with some dosimeters surpassing 20. The worst metrics appear in dosimeters with values closer to the extremes of the sRGB cube, like the green or the yellow uniform tone, whilst the grayer tones in the interior feature better ones. Overall, even if following a tendency of which reference values do better, variability amongst the improvement/worsening metrics indicates a different treatment of the colour values depending on their position in space.

When inspecting set 2 (white light and high brightness), the seen effect is the opposite. The predominant tendency is improvement of the quality when employing. Also, it can be observed that for the three calibrations, the general tendency is to worsen the metrics the more reduced the set of reference colours is. The linear, however, is more stable to reduction, whilst its improvement or worsening is more variable depending on the dosimeter. The most affected dosimeter is, again, the one that corresponds to the green tone, whilst the innermost show measures of the same proportion. The multi-dimensional calibration, again, shows the best qualities in a range between  $\Delta E$  2 and 10 in the huge majority of cases, and CCM shows distances of a more exaggerated error.

Sets 3 and 4 are shot under a high illuminant temperature (blue light), under high and low brightness each. The first thought that the observer might have is that the objective results might be parallel, with slight differences. In CCM, the evolution of the metrics for the experiments reducing the amount of values is similar, as they stay on the

TABLE 4. Accuracy CIEDE2000 ΔE metrics of the 64 colour patches of the self-made colour chart for the three colour calibration schemes in the four considered field images.

Table with 13 columns: Patch, Field Image 1 (Multi-dimensional, Linear, CCM), Field Image 2 (Multi-dimensional, Linear, CCM), Field Image 3 (Multi-dimensional, CCM), and Field Image 4 (Multi-dimensional, Linear, CCM). Rows contain ΔE values for patches 0-63.

same proportion of values. Nevertheless, set 3 shows more exaggerated differences with less colours, as it happens with the improvement percentages. The dosimeters that seem the less affected by the brightness difference between sets 3 and 4 are the three uniform brown ones, that show analogue improvements in both, whereas in the others the metric is more exaggerated. In the multi-dimensional calibration, this effect on the exaggeration of metrics in the images with a higher brightness is observed again. The metrics stay in the same range for both the uniform and textured dosimeters with the self-made chart, whereas the calibration effects with X-Rite’s seem to affect them differently. Nevertheless, X-Rite’s offers better results in general, with ΔE averagely staying between 2 and 10, except when involving a smaller set of reference colours. The linear calibration, as before, shows a big variability among the resulting metrics and improvements and worsening, revealing the effect of pixel discolouring.

The same information can be drawn from inspecting the CIEDE94 metrics, that follow a similar tendency, given that it features an alternate measure but in the same uniform colour space, CIELAB.

The same experiments have been ran over the four considered field images. However, considering that pictures of this kind are examples the subject of analysis of the whole system when deployed, the comparison between dosimeters has no place in this case. When testing over real field images, there do not exist any reference dosimeters, given that they are the subject of analysis of their colour change. Therefore, the accuracy of the calibration here is demonstrated by calculating the CIEDE2000 difference between the calibrated colour patches and the ideal hues only.

Table 4 shows the extracted CIEDE2000 values between the colour patches of every field image and the reference for the three calibration strategies. It is noticeable that, in spite



of being non-simulated test subjects, the objective results follow the same pattern than with the other sets of pictures. The Multi-dimensional calibration scheme offers the best results, ranging between CIEDE2000 0-1, whilst the linear stays between 0 and 10, and CCM exceeding that range by far, up to 40 in some cases.

The results for all experiments using different amounts of reference colours can be found in the supplemental table Field images. There it can be seen that, when executing the calibration via CCM, there exists a slight variability in the results in every different execution of the process. Meanwhile, the metrics for the linear and multi-dimensional schemes stay consistent for each colour for all experiments, meaning that the mapping is unique for every image with no possibility of miscalculation using the exact same reference values.

This is a proof that the conceived scheme is robust and performs well in the application context of this work.

Seeing all the valuable information given by the performed amount of experiments, the following results can be deduced:

- The usage of CCM for calibration is not recommended for monitoring with crowdsourced pictures, where pixelings in the acquired pictures and illumination shifts are a frequent thing to happen, introducing an extremely undesired variability among the results. It is especially important when considering to monitor similar content under different illuminations using the same intended reference. Furthermore, pixeling effects when acquiring the images are prone to happen when the chart is printed on paper, even if by professional means. This is a requirement in this project since charts are needed to be placed in many locations, and buying a big number of professional commercial charts would mean prohibitive costs. Easiness to reproduce is mandatory in this case.
- The multi-dimensional calibration approach does a better objectively-measured performance overall compared to the last linear proposal. Calculating the calibrated coordinates employing information from all three colour channels and not only one at a time helps reducing the pixel alteration that may affect the colour distributions when summoning an interpolation-based calibration scheme. Its metrics show that it builds the transfer functions accurately, mapping the anchor values correctly to the calibrated space and builds a solid interpolated region.
- In good lighting conditions, with a neutral temperature and high brightness, its measures indicate an excellent calibration quality. It is also robust to colour reduction in variable lighting conditions, which makes it an excellent candidate for calibration in museum and other indoors heritage exhibition room lighting setups.
- Comparing our choice of anchor colours with X-Rite's, it can be observed that the commercial chart selection does generally a good performance, whilst ours offers a more accurate experience at the right conditions. This arguably has its explanation in the physical presence of

the anchor colours in the real-life taken scenes. X-Rite's selection of colours is intended to reflect real, present hues (such as skin tones) that are likely to happen in real life, which means that they are contained in the interior areas of the sRGB cubes. That implies a closer distance between each other, and thus it forms a calibration hyper-surface with a finer assignment in the interior, less saturated regions. That also explains the bigger error in the dosimeters featuring colours in more exterior spatial positions as greens, and the different effect of illuminant temperatures on each hue. The conceived uniform sparsity of the sRGB space we designed, even after taking into account the effects of reality by taking the spectral measurements of the printed ideal colours, was produced with an idea of generality, instead of focusing in a finer calibration within a certain subset of the space. Nevertheless, its regular approach allows the reduction of the set of anchor colours without incurring in noticeable errors, and even improving the performance, in most of the cases. Given the cube-like distribution of the anchors, when removing the interior values whilst maintaining the corners the shape of the hyper-surface will remain approximately the same from one extreme to the other of the cube, effectively meaning a similar transfer function.

This structure is definitely important to consider when wanting to deploy the charts in real life scenarios. In order to reduce the invasiveness of the monitoring mechanism and its assets, which is a heavy factor to have in mind when designing similar systems, the size of the chart to be placed along the heritage piece should be as small as possible, especially the closer it is to the exposed item, for any reason. The size also influences the cost of the chart reproduction, so the smaller, the more possibilities of deployment it has.

In the end, our proposal of colour calibration chart offers more horizontal possibilities for the concerned kind of cultural heritage monitoring system. Whilst more sensitive to the lighting conditions than X-Rite's colour selection, it offers better results on cases of adequate brightness, up to the range of 11-36% of improvement in these situations, and fulfils the invasiveness, cost, reproducibility and robustness requirements expected from the application it is designed for.

## VI. CONCLUSION

In the previous text we have outlined a colour calibration technique for sRGB pictures gathered from commercially available cameras, and a colour calibration chart with a mathematically regular and generalist choice of colour values. Their performance and validity of usage for building a crowdsourcing-oriented colour calibration scheme for cultural heritage monitoring in controlled environments has been tested, their advantages stated, and compared with other alternative and widely used variants.

Whilst there already exist universal solutions for calibrating images, many of them already implemented in hardware



to use or in commercially purchasable or open-source software packages, concrete and punctual modifications of the same or different variants of these for specific applications may be needed to be implemented. However, sometimes these adaptations on them to match a particular use case might fail or not be suitable to be implemented, delivering results worse than what is expected.

In this paper, the aforementioned fact has been acknowledged. Available means, as the X-Rite chart and the CCM calibration, have been adapted to match to the projects requirements, and have been proved not being completely adequate to the conditions sought, failing to compel in some situations. Self-designed proposals have been conceived, and have been found to deliver more compelling results in these cases while being adaptive to the project from scratch.

In settings such as cultural heritage exhibition rooms, exposition conditions such as lighting can be a decisive factor, showing the boundary where custom solutions can be better. In this particular case of study, with high brightness and direct lighting.

The self-designed sRGB division in our chart offers equal or better results when reduced in anchors to be less invasive to the piece compared to the full range, and the multi-dimensional calibration scheme offers an accurate value mapping from the original sRGB to the calibrated space with many possibilities to be refined for concrete value ranges in further approaches.

Knowing the context of the application for a colour calibration system and its requirements, developing the adequate framework and assets is a feasible and reasonable possibility when the conditions of the sought aim do not allow the direct and bold implementation of available solutions due to any reason.

When monitoring any number of cultural heritage pieces in a controlled environment, it is desirable to incur in the least invasive mean possible. Crowdsourcing pictures implies not needing to install a camera system. The calibration chart should be easy and cheap to reproduce, possibly available in a big number if many pieces need to be put under surveillance, and reducible in size (colour patches) if less invasiveness is needed without meaning to worsen the performance. The calibration scheme should be automatic to perform and adaptive so it can be calculated from each picture without the need of a working hand, and easy to deploy without any specialised software while delivering satisfactory results.

The proposed scheme and chart are adequate for the followed aim in cases where other methods may fail, and are also cheap and easy to modify if future improvements may be needed.

The work presented here lays ground for further explorations in the direction of calibration chart design and specifically cultural heritage preservation. Knowing which range of the sRGB space the particular material to be monitored can reach in its values when measured, different proposals of anchor colours, encompassing a smaller region of the space around the values of interest, can be conceived. Therefore,

a finer calibration can be achieved in these regions, given that an effectively shorter distance between anchor points in space can be achieved. Aside, a mean for generating valid calibration charts depending on the size, amount of anchor colours and context of usage can be deduced for further improvement.

## ACKNOWLEDGMENT

The authors thank Ignacio García, from Once34, for his generous collaboration in the printing of the charts.

## REFERENCES

- [1] EC. *European Commission—Funding Opportunities for Cultural Heritage*. Accessed: Apr. 2021. [Online]. Available: <https://ec.europa.eu/culture/cultural-heritage/funding-opportunities-cultural-heritage>
- [2] S. Battiato, G. M. Farinella, F. Milotta, A. Ortis, L. Adesso, A. Casella, V. D'Amico, and G. Torrissi, "The social picture," in *Proc. ACM Int. Conf. Multimedia Retr.*, 2016, pp. 397–400.
- [3] H. Wirilander, "Preventive conservation: A key method to ensure cultural heritage authenticity and integrity in preservation process," *E-Conservation Mag.*, vol. 5, pp. 164–176, Oct. 2012.
- [4] L. Tomanidis, E. Bocaj, P. Kasnesis, and C. Z. Patrikakis, "Supporting cultural heritage preservation through game-based crowd-sourcing," in *Proc. 7th Strategic Innov. Marketing Tourism (ICSIMAT)*, Athenian Riviera, Greece, Oct. 2018, pp. 989–997.
- [5] A. D'Angelo, L. Ricciardi, and L. Gugliemetti, "Smart city and cultural heritage: Resilience through crowdsourcing involvement," in *Proc. IEEE Int. Conf. Environ. Electr. Eng.*, Jun. 2018, pp. 1–6.
- [6] 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.
- [7] M. Khalaf, "Smart cultural heritage: Technologies and applications," in *Proc. 2nd Smart Cities Symp. (SCS)*, 2019, pp. 1–6.
- [8] 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 *Built Heritage 2013 Monitoring Conservation Management*. Milan, Italy, Jan. 2013, pp. 737–744.
- [9] 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.
- [10] *Specification for Managing Environmental Conditions for Cultural Collections*, British Standards Institution, London, U.K., 2012.
- [11] W. Ma, E. Doehne, G. Bearman, M. Walton, and O. Cossairt, "Crowdsourced mobile phone images for heritage conservation monitoring," in *Proc. Digit. Heritage*, Oct. 2015, pp. 111–114.
- [12] 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.
- [13] XRite. *I1publish*. Accessed: Apr. 2021. [Online]. Available: <https://www.xrite.com/categories/calibration-profiling/i1publish>
- [14] ImageJ. *Fiji*. Accessed: Apr. 2021. [Online]. Available: <https://imagej.net/Fiji>
- [15] M. A. Barbero-Álvarez, J. M. Menendez, and J. A. Rodrigo, "An adaptive colour calibration for crowdsourced images in heritage preservation science," *IEEE Access*, vol. 8, pp. 185093–185111, 2020.
- [16] B. Ramírez Barat, E. Cano, M. T. Molina, M. A. Barbero-Álvarez, J. A. Rodrigo, and J. M. Menéndez, "Design and validation of tailored colour reference charts for monitoring cultural heritage degradation," *Heritage Sci.*, vol. 9, no. 1, Dec. 2021, Art. no. 41.
- [17] M. Haklay, *Citizen Science Volunteered Geographic Information: Overview Typology Participation*. Dordrecht, The Netherlands: Springer, 2013, pp. 105–122, doi: [10.1007/978-94-007-4587-2\\_7](https://doi.org/10.1007/978-94-007-4587-2_7).
- [18] 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, Oct. 2014.

- [19] C. Daffara, G. Marchioro, and D. Ambrosini, "Smartphone diagnostics for cultural heritage," *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 11058, Jul. 2019, Art. no. 110581K.
- [20] 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.
- [21] E. G. Hunt, *Study Museum Lighting Design*. San Marcos, TX, USA: Texas State Univ., 2009. [Online]. Available: <https://digital.library.txstate.edu/handle/10877/3203>
- [22] P. D. M. Fernández, F. A. G. Peña, T. I. Ren, and J. J. G. Leandro, "Fast and robust multiple ColorChecker detection using deep convolutional neural networks," *Image Vis. Comput.*, vol. 81, pp. 15–24, Jan. 2019.
- [23] C.-L. Chen and S.-H. Lin, "Intelligent color temperature estimation using fuzzy neural network with application to automatic white balance," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, Oct. 2010, pp. 796–803.
- [24] S. Sunoj, C. Igathinathane, N. Saliendra, J. Hendrickson, and D. Archer, "Color calibration of digital images for agriculture and other applications," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, pp. 221–234, Dec. 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924271618302600>
- [25] M. Ashraf and L. R. Sapaico, "Evaluation of color correction methods for printed surfaces," in *Proc. Colour Vis. Comput. Symp. (CVCS)*, Sep. 2018, pp. 1–6.
- [26] XRite. *Colorimetric values for ColorChecker Family Targets*. Accessed: Apr. 2021. [Online]. Available: <https://xritephoto.com/colorcheckerkits>.
- [27] Imatest. *Imatest Products*. Accessed: Apr. 2021. [Online]. Available: <https://www.imatest.com/products/>
- [28] Imatest. *Color Correction Matrix (CCM)*. Accessed: Apr. 2021. [Online]. Available: <https://www.imatest.com/docs/colormatrix/>
- [29] J. Vaillant, A. Clouet, and D. Alleysson, "Color correction matrix for sparse RGB-W image sensor without IR cutoff filter," in *Unconventional Optical Imaging*. Strasbourg, France: SPIE, Apr. 2018, p. 3. [Online]. Available: <https://hal-cea.archives-ouvertes.fr/cea-02390508>
- [30] A. Abdalla, H. Cen, E. Abdel-Rahman, L. Wan, and Y. He, "Color calibration of proximal sensing RGB images of oilseed rape canopy via deep learning combined with K-means algorithm," *Remote Sens.*, vol. 11, no. 24, p. 3001, Dec. 2019.
- [31] Y. Motonaga, T. Matsumoto, and N. Motonaga, "Color chart of European Pear 'Le Lectier' based on the color image analysis," in *Proc. SICE Annu. Conf.*, 2010, pp. 2455–2461.
- [32] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sep./Oct. 2001.
- [33] *Technical Report CIE 199: 2011 Methods for Evaluating Colour Differences in Images*, CIE, Cambridge, U.K., 2011.
- [34] B. Lindbloom. *Delta E CIE 1994*. Accessed: Apr. 2021. [Online]. Available: [http://www.bruce.lindbloom.com/index.html?Eqn\\_DeltaE\\_CIE94.html](http://www.bruce.lindbloom.com/index.html?Eqn_DeltaE_CIE94.html)



#### MIGUEL ANTONIO BARBERO-ÁLVAREZ

received the bachelor's degree from the Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, ETSI Telecomunicación, Universidad Politécnica de Madrid, in 2016, with a thesis about synthesizing contents for UHD television via super-resolution, and the Máster Universitario en Ingeniería de Telecomunicación from the Universidad Politécnica de Madrid, in 2019, with a 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 at TensorFlow and Keras for emotion recognition and detection of sleep apnea out of speech data with GAPS Group, Universidad Politécnica de Madrid. 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's specialization.



**JUAN ANTONIO RODRIGO** received the degree (Hons.) in telecom engineering, 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 has been a member of the Visual Application Telecommunication Group (GATV), since 2007. During these years, his work has been focused in 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. He has several international congresses and publications endorse these research and development deployments.



#### JOSÉ MANUEL MENÉNDEZ

(Senior Member, IEEE) was the Director of the Visual Telecommunication Application Research Group—GATV, from 2004 to 2019, and has been the Chair of the Spanish Public Broadcaster RTVE, UPM, since 2015. He is currently a Full Professor at the Signal, Systems and Radio-communications Department, ETS Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, 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 papers 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. He has been a Regular Reviewer of the IEEE Signal Processing Society, since 2000, for different international journals and conferences sponsored by that entity. He has been 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.

• • •