Exploring the Limits of Color Accuracy in Technical Photography

Eric Kirchner (eric.kirchner@akzonobel.com)
AkzoNobel Paints and Coatings, Color Technology, Rijksstraatweg 31, 2171 AJ Sassenheim, The Netherlands
https://orcid.org/0000-0002-9379-8802

Carola van Wijk
Rijksmuseum

Henni van Beek
Rijksmuseum

Tammo Koster
Akzo Nobel NV

Research article

Keywords: color management, digitization, Cultural heritage recording, camera profiling

Posted Date: February 8th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-162431/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

The growing importance of publishing art collections online has led to increasingly strict tolerances on digital photography of art objects. There are two internationally recognized sets of guidelines for creating high-quality digital images, Metamorfoze and FADGI.

These guidelines require using sets of standardized color patches in museal photography. The X-Rite ColorChecker SG chart with 140 color patches is often used. Recent studies showed that even in the standardized conditions of museum photography studios it is often very difficult to satisfy the strictest guidelines on color accuracy for camera profiling, with no indications for improvements.

Here we report results of our investigation into the bottlenecks in achieving high color accuracy. We show that a large part of the color deviations originates from the 15 black color patches of the ColorChecker SG chart. These patches have a large impact on the average color deviation and the maximum color deviation that are the performance measures for color accuracy in the FADGI and Metamorfoze guidelines. We show that spectrophotometer measurements for the black patches produce color deviations \(dE(\text{CIE 1976})\) ranging from 3.7 to 5.2 with respect to reference data, making it impossible to meet the strictest Metamorfoze guidelines. The black patches push the average color difference CIEDE2000 from 0.59 to 0.82 with respect to reference data already when using spectrophotometer data. Since the strictest FADGI guidelines prescribe an average CIEDE2000=2.0, this leaves little tolerance for errors due to lighting and camera profile.

Our results indicate that the common practice of manually tweaking camera profiles until the software suggests sufficient color accuracy is obtained with respect to suppliers’ reference data often does not improve color representation but makes it worse, especially for representing dark nuances. This is unfortunate for example for the digital photography of 17th century Dutch paintings, where dark passages occupy large areas of the art works.

We show that the key step in achieving color accurate digital photography is to use customer reference data rather than commonly used generic reference data. We explain the results by investigating not only the color properties of the ColorChecker SG chart, but also its glossiness.

1 Introduction

Many major museums and art institutes run large multi-year projects to digitize their collections for decades now [1]. Since the aim of the digitization process is to obtain as close a digital representation of the original object as possible [2], the collected data needs to be of the highest accuracy and quality [3]. Digital photography and color management have been recognized as important components of this process. Therefore photograph studios of museums are continuously developing workflows and standardized imaging systems in which color management plays a major role.
Also in the technical analysis of cultural heritage the role of imaging science is growing. Several trends combine to lead to increasingly tighter color tolerances for technical photography. For example, photographs are increasingly used to help making decisions on conservation strategies, making it essential that conservators can trust the accuracy of the details they see in these images. Photographs are also used more than before to evaluate the progress and results of conservation campaigns. In this case, photographs captured before, during and after the conservation treatment need to be compared with each other, which requires being able to exactly reproduce the conditions and camera profiling after months or years. A third example is the growing awareness that for investigating art works in relation to pieces located in other museums it is essential to use highly reproducible and repeatable images.

The development of processes for technical photography at color tolerance levels tighter than ever before is also an enabler for new approaches and technical analyses. A first striking example of this is the extremely high resolution photography of historical paintings which is being developed over the past few years, which enables new technical analyses and conservation approaches [4][5]. High resolution imaging requires composing increasingly larger numbers of partial photographs into one overall image. For example, for the photography of Rembrandt's *Marten* and *Opjen* paintings at 1250 ppi resolution, 242 partial photographs (also known as image tiles) had to be stitched and registered in 2016 [4], while for the most recent photography of Rembrandt's *Night Watch* a total of 11845 image tiles were captured and combined [6]. In order to ensure color consistency throughout the composite image, and to make automated stitching of image tiles feasible, very tight color tolerances are required within each image tile.

A second example of the benefits of using highly standardized processes for capturing photographs of artworks is that parts of this work can be outsourced to external institutes or companies. The Rijksmuseum is already working for years on digitizing their collection of one million art objects with high resolution photographs and making this publicly available [1]. For achieving this goal it is essential to be able to outsource art of the work to external contractors that are able to work according to strict guidelines.

However, various investigations have shown that even when working with professional cameras and well-controlled lighting set up, it is often very difficult to satisfy the highest color quality level as defined by internationally recognized guidelines [7][8][9].

In this article we will identify the cause of a large contribution to the color deviations that create these difficulties. Section 2 summarizes the current process for digital photography, including camera profiling, when working according to internationally recognized guidelines. Section 3 introduces the X-Rite ColorChecker SG chart, which is most often used for camera profiling.

In Sect. 4 we will show how the black color patches in color charts are a major contributor to color deviations in the camera profiling process, and we will analyze why this is the case. We will show how color quality improves by avoiding the use of generic ground truth data. Improved color quality can be obtained by introducing custom spectrophotometer measurements, and does not require using highly expensive instruments.
We summarize our conclusions in Sect. 5, and identify the most promising routes for further improvements in color accuracy.

## 2 Methods

This study aims to identify the main contributions to color deviations in the technical photography process when using current guidelines. Over the past decades the science behind accurate color reproduction has resulted into consensus for most aspects of the process. An early summary was written by Berns [10]. Currently there are two internationally recognized sets of guidelines to standardize the work flow for creating high-quality digital images of cultural heritage. FADGI [11] was originally developed by the Library of Congress in the US, whereas Metamorfoze [12] was developed by the National Library in the Netherlands. These guidelines cover more aspects than only color, such as spatial resolution and misregistration. In the present work we will only discuss quality aspects related to color.

The FADGI and Metamorfoze guidelines are very similar in many aspects, and both aim at satisfying the same set of ISO standards for measuring imaging performance. For the large majority of metrics, they agree on aim and tolerance levels [13]. For the performance measure for color tolerance Metamorfoze chooses dE(CIE 1976), whereas this is CIEDE2000 for FADGI (both measures defined by the International Commission on Illumination (CIE) [14]).

After pioneering work by Martínez et al. [15] to develop a procedure to optimize the reproduction process, different workflows have emerged for creating images according to these guidelines. Since there is no consensus on the best workflow, we limit ourselves here to sketching a procedure that includes the most essential steps as used by technical photographers in the Rijksmuseum, as well as described in recent publications [16][2][17].

1. Flat-field calibration. The camera aperture is varied to optimize sharpness. A uniform chart is captured, and the standard deviation in brightness across the image is minimized. This compensates for nonuniformity in lighting as well as deviations caused by the camera lens [18].
2. White calibration. This is to ensure that the white color is captured correctly. The camera exposure time and/or lighting power are optimized by comparing camera RGB values with the known CIE-Lab value of a white standard, consistent with the D50 white point.
3. Fine-tune brightness variation across the image. The X-Rite ColorChecker SG chart is captured with the camera. Using only the white color patches of this chart, a final verification is obtained to ensure that brightness variation is small across the image.
4. Calibration of the Tone Response Curve (TRC), also known as the Opto-Electronic Conversion Function (OECF). Using all color-neutral patches of the ColorChecker SG chart, the goal of this step is to make sure that all neutral colors between white and black are mapped to accurate CIE-L* values. The color deviations (or color encoding errors, as they are called in FADGI) need to be within the tolerances provided by the guidelines. This procedure accounts for the nonlinearity of the camera system. It ensures contrasts in the image correspond to visually perceived contrasts. By making sure
that the a* and b* values that result for the achromatic patches are sufficiently close to zero, the white balance is checked across the image and a color cast is avoided.

5. Verify sampling efficiency. Evaluate what percentage of intended image resolution (in ppi) is obtained.

6. Camera characterization, also known as building the camera (ICC) profile. The full X-Rite ColorChecker SG chart is captured. Software detects all color patches in the image and calculates average RGB values. The camera (ICC) profile is built by optimizing model parameters in such a way that converting the RGB values of each patch leads to CIE-Lab values with a minimized color difference with respect to reference CIE-Lab data for the patch.

7. Determine image quality (see below).

We note that the camera (ICC) profile thus created is only valid for the camera under the conditions used when capturing the images with the color patches. It is assumed that the same ICC profile remains valid for the photography of other objects, as long as the lighting and camera conditions (such as lenses) remain unchanged. The custom ICC profile is assigned to the photographic image, replacing the standard color ICC profile color space [2].

In the last step of the process, the image quality is assessed. It would have been best if this would include a visual assessment that quantifies the quality of the image in terms of its accuracy in reproducing the original art object. However, there is no standard procedure for this and it would be highly subjective. We will not elaborate on this aspect further, other than by remarking that during such a subjective assessment the art object should be lighted with the same light that was assumed in the camera profiling (i.e., D50 spectrum), and the image would need to be visualized on a color calibrated monitor display (page 231 of [19]).

Instead of a visual assessment of image quality, the color accuracy of the process is expressed in the color deviations (or color encoding errors, as they are called in FADGI) that are tolerated in these guidelines. The quality of the image is evaluated by determining color differences for all 140 color patches. The color difference for each patch is calculated between on the one hand CIE-Lab values calculated with the camera profile using the camera RGB values as input, and on the other hand reference CIE-Lab data as measured by a spectrophotometer.

The strictest FADGI guidelines require the average to be smaller than CIEDE2000 ≤ 2 [20]. We note that the previous version of FADGI demanded a CIEDE2000 of 3.0 on average, and a maximum value lower than 6.0 [11].

For Metamorfoze, the strictest guidelines demand the maximum color difference to be dE(CIE 1976) ≤ 10.0 and the average to be dE(CIE 1976) ≤ 4.0 [12]. Apart from that, for neutral patches it is required that dL* ≤ 2, dC* ≤ 2 and dE(CIE 1976) ≤ 2.83 [12].

3 Color Charts
The previous section shows that most steps in the process of camera profiling require charts with series of standardized color patches. Here we will introduce two color charts that we used in our investigation: the X-Rite Digital ColorChecker SG chart (Figure 1) and the Munsell Linear Gray Scale chart (Figure 2). We note other charts are available as well: the Golden Thread Object-level Target and Device-level Target [8][21][22], the Universal Test Target and the Next Generation Target [23][24][25].

3.1 ColorChecker

A) Original ColorChecker

The current X-Rite Digital ColorChecker SG is a direct descendent of the original MacBeth ColorChecker Chart. Under the full name of Macbeth ColorChecker Color Rendition Chart [26] it was introduced in 1976, and it is still for sale [14][27]. Its color selection is based on scientific work by McCamy et al. in 1976, representing frequently occurring colors in common photography. The colors include skin tones, foliage, flowers and the blue sky, but also the primary colors of additive and subtractive mixing [28].

This chart contains 24 color patches on a 21.6 x 29.2 cm cardboard, and all colors are matte. Measuring 40x40 mm, the color patches are larger than those on other color charts. Earlier research showed that the color of each patch is not completely homogeneous, with variations up to 1 or 2 units in CIE-Lab coordinates [29]. These variations are expected to be much smaller for the other commercial color charts, since they have much smaller sized color patches.

The MacBeth ColorChecker was not intended to create camera profiles: digital cameras were yet to come when it was introduced. Instead, it aimed at being included as a useful color reference in high-end photography of objects [30]. It soon indeed became the *de facto* standard in high end photography [31], and even today it remains to be so in professional photography under the name X-Rite ColorChecker Classic [14]. When digital cameras became popular it was found that the MacBeth ColorChecker was less suitable for this task, as the colors of this chart cover only a relatively small color gamut [30].

B) ColorChecker DC

The Gretag-Macbeth ColorChecker DC was launched in 2000 specifically for profiling Digital Cameras (DC). For this reason the number of colors and also the color gamut covered by them was greatly expanded: the new chart contained 237 color patches [26]. Eight of these patches were high-gloss, and all others were matte [32]. The high-gloss patches were probably included to cover a wider color gamut: the high gloss colors showed a higher color saturation than what is possible with matte colors. However, in practical use the high-gloss patches often picked up reflections, creating out-of-gamut colors in the image that ruined the resulting profile [30]. Also, the camera profiles created with the Gretag-Macbeth ColorChecker DC were often not smooth enough, possibly due to the very large number of color patches in this chart [33]. For these reasons this chart was soon replaced by the ColorChecker SG.
C) ColorChecker SG

In 2005, GretagMacBeth launched the Digital ColorChecker SG, just before the company was taken up by X-Rite [34]. The new chart contains an extended color palette of 140 patches, all of them being Semi-Gloss (SG). Since the patches are more glossy than those on the earlier ColorChecker charts, more saturated colors are reached, and also darker colors are reached by the Digital ColorChecker SG chart. Compared to the original MacBeth ColorChecker Chart it provides a wider variety of skin tones, and a series of achromatic colors to determine the camera white balance under different light sources [14]. Just like the simpler MacBeth ColorChecker Chart, the Digital ColorChecker SG chart is widely used by professional photographers. For the most demanding applications, such as the photographic documentation in the conservation of cultural heritage, the ColorChecker SG replaced the MacBeth ColorChecker Chart as the common color chart [35]. In order to enable this group of highly professional photographers to build camera profiles and to do camera characterization (cf. step 6 in the previous section), X-Rite produced reference data obtained with a spectrophotometer for each of the 140 color patches, and made this data publicly available [36].

In 2015 X-Rite announced that “for regulatory and compliance reasons”, the color formulations of all color patches in the X-Rite ColorChecker Classic and the X-Rite Digital ColorChecker SG had to be changed. This meant that for both charts there are different versions of the colorimetric ground-truth data, suitable for charts produced before or after November 2014 [36].

3.2 Munsell Grayscale

For assessing image quality, it would have been best to use a different set of standardized color patches than the set used for building the camera profile. However, by lack of materials in practice often the X-Rite ColorChecker SG chart is used for both steps in the process. In order to create some independent data at least for the neutral colors sometimes the Munsell Grayscale is used.

The Munsell Linear Gray Scale [37] contains 18 semi-gloss neutral grey color patches, with colors equally spaced from almost white (L*=95) to almost black (L*=10). It also contains two high-gloss black patches with L*=5, making it darker than any patch on the X-Rite ColorChecker SG chart.

4 Results And Discussion

4.1 The nature of the color deviations and their magnitude

Based on many years of experience as professional photographers for the Rijksmuseum, two of the present authors (CW and HB) confirm the conclusion from several other investigations [7][8][9]. It is very hard to ensure that in step 7 of the workflow of digital photography (cf. section 2) the color quality and contrast of the image is evaluated to be meeting the highest quality level as defined by the internationally recognized guidelines. We reached this conclusion even when working with high-end professional cameras (Hasselblad) and well-controlled high-end professional lighting set up (Broncolor).
We used software package iQ-Analyzer [38] to analyze the image quality. We also used the calculational tool RM that was developed by the Rijksmuseum (co-author HB), which can analyze Metamorfoze, FADGI, X-Rite standards and customer look up tables. As mentioned in Sect. 2, the image quality is quantified by calculating for each color patch the color difference between on the one hand CIE-Lab color coordinates calculated from camera RGB values using the camera (ICC) profile, and on the other hand reference CIE-Lab data published by X-Rite and created with a spectrophotometer [36].

All these analyses indicate that the largest deviations are found for the darkest color patches of the X-Rite ColorChecker SG chart. In order to further analyze the origin of these color deviations, we decided to measure color coordinates for all 140 color patches with a spectrophotometer ourselves. For this we used an i1Pro instrument (supplier: X-Rite), and in order to avoid measurement errors we took the average reading from three consecutive color measurements for each patch. We used this custom data to calculate the color difference with the supplier's reference data [36] for each patch in the chart. All calculations were done in terms of CIEDE2000, with D50 illuminant and CIE 2° standard observer.

The results are shown in Fig. 3. Colors indicate the magnitude of the color difference. It is clear from Fig. 3 that our custom data shows the largest deviations from the reference data for the black patches. This is the case for all black patches, both those dispersed along the circumference of the chart as well as patch E6. Based on many years of experience as professional photographers for the Rijksmuseum, two of the present co-authors (HB and CW) confirm that this pattern is typically found also when evaluating image quality after building camera profiles. The black patches almost always give large deviations in the value for lightness L*, producing color differences large enough to make it more difficult to satisfy the strictest criteria of FADGI and Metamorfoze guidelines [39].

In terms of CIEDE2000, the average color deviation for all 140 color patches is 0.82. For the 15 black color patches the average color deviation is CIEDE2000 = 2.74. If the black patches are ignored, the average color deviation of the remaining 125 color patches is CIEDE2000 = 0.59. For camera profiling, the strictest FADGI guidelines require the average color deviation to be smaller than CIEDE2000 ≤ 2 [20]. Since the quality achieved in camera profiling depends also on errors due to lighting and camera optics, it is concerning that using different sets of spectrophotometer data already shows errors of this magnitude.

In terms of dE(CIE 1976), the black patches show color deviations of dE(CIE 1976) = 4.23 on average (maximum: 5.23), increasing the average for all patches from dE(CIE 1976) = 0.88 to 1.24. The strictest guidelines for camera profiling using the Metamorfoze guidelines demand for each of the neutral patches that dE(CIE 1976) ≤ 2.83 [12]. This is clearly not fulfilled when comparing data from different spectrophotometers, making it impossible that camera profiles will satisfy this demand. There is no problem yet with the Metamorfoze guideline that the maximum color difference needs to be dE(CIE 1976) ≤ 10.0 and the average for all 140 patches should be dE(CIE 1976) ≤ 4.0 [12]. However, with the large color deviations reported here when comparing data from spectrophotometers there is hardly any tolerance possible for errors due to lighting and camera optics.
The large color deviations (mainly in L*) for black color patches is exemplified by black patch E6. According to X-Rite's reference data [36], its color coordinates (L*, a*, b*) are (8.07, 0.12, -0.93), whereas our spectrophotometer data for the same patch show (12.75, 0.22, 0.07). The color difference between both sets of color coordinates is CIEDE2000 = 3.11 and dE(CIE 1976) = 4.79. Our measurement data show a much larger value especially for lightness coordinate L* = 12.75 as compared to the value L* = 8.07 provided by X-Rite. The lightness difference strongly dominates the overall color difference for this patch. Also for all other black patches our measurement data show considerably larger lightness values than the corresponding reference data. We note that in the OpenDICE software, for patch E6 a value of L*=10.88 is used, based on measurements with a Konica Minolta FD-7 (45°/0°) spectrophotometer [40]. This confirms our conclusion that the lightness values of the black color patches may be too small in X-Rite's publicly available reference data.

Our spectrophotometer data are based on measurements on a relatively new X-Rite ColorChecker SG chart. At the back side of the chart the production date is printed as being June 2018. To make sure that our results are not the result of a bad copy of the chart, we repeated our measurements with an older X-Rite ColorChecker SG chart, with production date July 2016. Also in this case we find a lightness value that is much larger than in X-Rite's reference data.

We thus find that our measurement data differ from X-Rite reference data especially by yielding much larger values for L* for the dark color patches. As long as this problem is not solved, photographers are faced with the following dilemma when building camera (ICC) profiles.

If photographers would choose to use their own measurement data instead of X-Rite reference data, the darkest colors in calibrated images would not be pitch black, but they would probably correspond to values of L* ≥ 12.75. This option would effectively make the darkest spots in calibrated photographs brighter than absolute black. By using the camera profiles created in this way for the photography of artworks, many subtle nuances in dark colors on the original art work will be lost in the digital reproduction. This is unfortunate, since dark nuances form an important aspect of many art works, such as seventeenth century Dutch paintings.

Alternatively, a photographer may choose not to use own measurement data for the black patches. In that case, in order to be consistent, the photographer should then also not use own measurement data for the other color patches. Choosing this alternative option would also be unfortunate, because the photographer would have to assume that all color patches in an individual ColorChecker SG chart are well described by the reference data. In practical situations, photographers may want to use a ColorChecker SG chart that is not brand new, so its colors may have slightly changed in time due to color fading. If high quality camera profiles can only be built by using new ColorChecker SG charts, this would become a costly affair.

4.2 The role of instruments in creating the color deviations
In the previous section we found that our measurement data on the X-Rite ColorChecker SG chart differ from X-Rite reference data especially by producing much larger values for $L^*$ for the dark color patches. One possible explanation for this problem may be that these datasets were not produced by exactly the same instrument. Our reflectance measurement data was taken with an i1Pro instrument, whereas X-Rite's reference data was produced with an i1Pro2 instrument [36]. However, these two instruments are sufficiently similar to expect that they are well aligned with each other. Both instruments have the same measurement aperture, with an illumination spot size of 3.5 mm [41][42]. Both instruments are based on the same $45^\circ/0^\circ$ measurement geometry (with ring illumination), which should make us expect their results to be aligned by a color difference of $dE(CIE 1976)$ between 0.4 and 0.8 on average [35]. As reported in the last paragraphs, the color deviations for which we are trying to find an explanation here are much larger than this value.

However, when we discussed these issues with an instrument supplier in 2020, it was suggested that the problem could be solved by not using an i1Pro instrument but a much more expensive spectrophotometer. Obviously, we should choose an instrument with the same $45^\circ/0^\circ$ measurement geometry as the instrument that was used to create the reference data. For example, it is well known that using a diffuse sphere $d/8^\circ$ instrument instead of a $45^\circ/0^\circ$ instrument would lead to large differences in reflectance data, the magnitude of the difference depending on the glossiness level of the sample [43][44].

We considered the hypothesis that instrumental inaccuracy may be larger in our investigations than in previous investigations, because we are focusing on very dark colors. For spectrophotometers black colors may be the most difficult to obtain a good color accuracy, since they lead to a relatively low signal-to-noise ratio. We investigated if inaccuracies of the i1Pro instrument could have caused the measured color deviations.

We decided to verify the accuracy of the i1Pro with a five times more expensive high-end spectrophotometer, the spectro2guide from supplier BYK-Gardner. Like the i1Pro, the spectro2guide is a $45^\circ/0^\circ$ instrument, and also with ring illumination. It has the additional advantages of being able to also measure the gloss-level of a paint sample at $60^\circ$, and to detect fluorescence.

It is very challenging to measure the color patches of the X-Rite ColorChecker SG chart directly with a spectro2guide, since its measuring area with a diameter of 8 mm [45] is almost exactly the size of a color patch. For this reason, we first did a series of measurements on relevant paint samples with sizes large enough to enable care-free measurements.

We collected a range of samples with neutral colors, ranging from dark to pitch black as well as some neutral samples with brighter neutral colors. The samples also varied in gloss level as indicated in the following list:

- 7 samples with presumably the same black color S-9000 N, while varying in gloss level from very matte up to high gloss. These samples are the blackest samples of the NCS Gloss Scale (supplier:
Natural Color System) [46].

- 21 color patches from the Munsell Linear Gray Scale [37]. This set contains two high-gloss very black patches, and 18 semigloss neutral grey color patches ranging from very dark to very light.
- 7 car refinishes paint samples with the same black basecoat, covered by clearcoats (Mix & Matte, supplier AkzoNobel) that vary from matte to very high gloss.

The measurement data for these 35 samples are summarized in Table 1. The data shows that this set of samples provides a nice selection of dark samples, many with $L^*$ values lower than 10, while also showing a variety of gloss values (at 60°) ranging between 2.0 and 93.2 Gloss Units. Therefore, this set should enable a detailed comparison between the accuracy of the i1Pro and spectro2guide instruments especially for dark samples at various gloss levels.

| Samples                              | $L^*$ (i1Pro) | $L^*$ (spectro2guide) | Gloss values (spectro2guide) |
|--------------------------------------|--------------|----------------------|-----------------------------|
| 7 NCS Gloss Scale S-9000N            | 5.4–23.6     | 5.1–22.9             | 2.0–93.2 GU                 |
| 2 darkest patches                    | 4.4–5.1      | 4.3–4.6              | 79.9–81.9 GU                |
| 19 other patches from the Munsell Linear Gray Scale | 9.0–93.6 | 8.4–92.6 | 13.2–21.9 GU |
| 7 Car Refinishes Mix & Matte (AkzoNobel) | 0.8–5.0   | 0.5–4.3              | 10.2–89.7 GU                |

The data of Table 1 are plotted in Fig. 4, and they show that the lightness value $L^*$ as measured by both instruments agree very well with each other. For almost all samples, the i1Pro measures an $L^*$ value that is approximately 0.6 units lighter than the value measured with the spectro2guide. Figure 4b makes it clear that the difference in $L^*$ values as measured by the two instruments does not vary systematically with lightness $L^*$.

On the other hand, Fig. 4b does show that atypical values for this difference in measured $L^*$ values are found for the Car Refinishes Mix & Matte samples. The reason for this becomes more clear in Fig. 5, where we show the same difference in $L^*$ values as measured by the two instruments, in this case plotted as function of measured gloss value at 60°. This shows that the Car Refinishes Mix & Matte samples do follow the same trend as for the other samples: lightness values measured by the i1Pro are generally 0.6 units lighter than the value measured with the spectro2guide, except when the measured sample has a gloss value lower than 25 gloss units. For those very matte samples, the difference may increase up to values of 1.0 or even 1.6.
These results show that the i1Pro instrument does not cause the color deviations for black patch E6 of the X-Rite ColorChecker SG chart mentioned in Sect. 4.1. We noticed that the i1Pro measured $L^* = 12.75$ where X-Rite reference data mentions $L^* = 8.07$, i.e. a difference of $dL^* = 4.68$. Figure 5 shows that also for extremely matte colors the i1Pro is much more accurate than this. Even if patch E6 would have lost all its glossiness in time, reaching zero gloss units, this would not explain the measured color differences.

**4.3 The role of the chart in creating the color deviations**

By eliminating the possibility that instrumental inaccuracies cause the color deviations reported in Sect. 4.1, they are most likely explained by variations in the X-Rite ColorChecker SG chart itself.

However, since we find similar color deviations for charts produced at several moments in time (July 2016, June 2018, July 2019), it is unlikely that production variation can explain the measured color deviations. In that case, we would expect more variability in the color deviations. And since we find the color deviations not only charts that are 4 years old, but also for charts that are less than 1 year old, it is also unlikely that the color deviations are due to ageing of the charts or fading of its colors.

When we discussed our results with X-Rite, they were very cooperative in our search for the origin of the color deviations. X-Rite provided one of the original X-Rite ColorChecker SG charts that they had used in 2014–2015 for creating the reference data on their website [36]. With the i1Pro instrument we measured the black patches of this old X-Rite ColorChecker SG chart from 2014. This resulted in values for lightness $L^*$ varying between 8.0 and 8.87, as shown in Table 2. Only black patch C10 seems to deviate by having $L^* = 9.68$. For patch C10 repeating the measurement with the i1Pro and also verification with the spectro2guide instrument showed consistently larger $L^*$ values than for the other black patches (as shown in Table 3). These larger $L^*$ values for patch C10 may be caused by (invisible) scratches, dust and small paper particles clustering on glue residues near the edges of the color patch, ageing due to being exposed to light etc.

For patch E6 we find $L^* = 8.38$ with the i1Pro. We conclude that the reference data on X-Rite's website is correct for the original X-Rite ColorChecker SG chart from 2014, and that this data is correctly reproduced by the i1Pro instrument.

We also conclude that the color deviations for which we are seeking a solution in this article are caused by differences in color and/or glossiness in X-Rite ColorChecker SG charts produced in July 2016, June 2018 and July 2019 as compared to the version from 2014.

We investigated this further by measuring reflectance values, lightness values $L^*$ and glossiness values at 60° geometry using the spectro2guide instrument. For all black patches and all wavelengths, the reflectance values on the X-Rite ColorChecker SG chart produced in July 2019 are larger than the corresponding values on the original chart from 2014. On the original chart, all black patches have a reflectance value of 0.9% for all wavelengths. On the chart from July 2019, the reflectance is larger by 0.6% for small wavelengths, with the difference dropping to 0.1% for large wavelengths. This makes the black patches on the new chart not only lighter but also slightly more bluish than on the chart from 2014.
The difference in reflectance values between the old and the new chart may be caused by differences in glossiness. This was confirmed by measuring the Gloss value at 60° geometry using the spectro2guide instrument. Table 4 shows that for the original chart from 2014, the black patches have a gloss value that ranges from 25.7–27.0 Gloss Units. For the chart from July 2019, the corresponding variation is between 17.7 and 19.2 Gloss Units. These data show that in comparison to the chart from 2014, the black patches have increased in L* values by 1.7–2.9 units, while decreasing in glossiness by 7.0–9.1 Gloss Units.

| L*       | Black patches       | White patches    |
|----------|---------------------|------------------|
| Chart version 2014 | 8.0–8.87; 9.68 (E6: 8.38) | 96.22–96.51 |
|          | (C10: 9.68)         |                  |

| L*       | Black patches       | White patches    |
|----------|---------------------|------------------|
| Chart version 2014 | 7.82–8.52; 8.98 (E6: 7.94) | 94.72–94.93 |
|          | (C10: 8.98)         |                  |
| Chart version 2019 | 9.74–10.79; 10.85 (E6: 10.14) | 95.07–95.51 |
|          | (C10: 10.85)        |                  |

| Gloss    | Black patches       | White patches    |
|----------|---------------------|------------------|
| Chart version 2014 | 25.7–27.0 Gloss Units | 20.7–21.7 Gloss Units |
| Chart version 2019 | 17.7–19.2 Gloss Units | 17.3–19.2 Gloss Units |

We performed a similar analysis on the white patches of both charts. The range of L* values changed from 94.72–94.93 on the chart from 2014 to 95.07–95.51 on the chart from July 2019. For the glossiness, we find values ranging from 20.7–21.7 Gloss Units for the original chart and from 17.3–19.2 Gloss Units for the chart from 2019.
We conclude that in the chart from 2019 the glossiness of the black patches aligns well with those of the white patches. In the 2014 version of the chart the black and white patches were not aligned with each other in terms of gloss values. Possibly in 2014 the black patches were created with a relatively high glossiness, in order to achieve very dark L* values. We already saw in Sect. 3.1 that reaching darker blacks and obtaining a wider color gamut was one of the objectives when X-Rite introduced the ColorChecker SG chart. In later versions of the chart the glossiness of the black patches may have been made to align closer to those of the white patches, resulting in slightly less dark patches.

Conclusions

We have investigated the feasibility of two sets of internationally recognized guidelines for technical photography, FADGI and Metamorfoze, as far as color accuracy is concerned. In the process of camera profiling, a color accuracy of 2.0 (average) and 6.0 (maximum) is needed when following the highest level FADGI guidelines (and equivalent tolerances for Metamorfoze) [20]. Several earlier investigations showed that even when working in laboratory conditions it is difficult to meet the best color image quality level of these guidelines [7][8]. In this article, we investigate why even professional photographers, using high-end cameras and well standardized lighting conditions, often obtain camera profiles that produce color deviations that are too large to satisfy the guidelines. Only by manual tweaking of the camera profile these color deviations can sometimes be reduced to acceptable levels, but this makes the whole process time consuming and badly reproducible.

Our results show that these problems arise predominantly for the black color patches of the X-Rite ColorChecker SG chart, which is the most commonly used target for professional photography. When we measured all color patches this chart with a spectrophotometer, our data indicated that relatively large color deviations occur already when our spectrophotometer data are compared to the spectrophotometer reference data published by X-Rite. Indeed, these color deviations occur predominantly for the black color patches, and their magnitude is large enough to be a major contributor for camera profiles to fail meeting the strictest guidelines.

Although our measurement data were created with an i1Pro, which is not a high-end spectrophotometer, we showed that the color deviations would not be solved by using a much more expensive spectrophotometer. Instead, our results show that the color deviations are caused by the X-Rite ColorChecker SG chart itself. We investigated copies of the chart that had been produced in July 2016, June 2018 and July 2019. We showed that all of them differ significantly from the version produced in 2014–2015, which was used to produce the reference data published on X-Rite’s website. The charts from 2016 and later show considerably higher values for lightness L* and lower glossiness values than the chart from 2014–2015. This completely explains the color deviations that we found when comparing our own spectrophotometer data with the reference data.

The later versions of the ColorChecker SG chart have changed over the years since 2014. This makes it more difficult to satisfy the strictest tolerances implied by guidelines such as Metamorfoze and FADGI, if
the X-Rite reference data from 2014 are used as a reference. In the Metamorfoze guidelines [12], for each neutral patch (which includes the black patches investigated here) the strictest maximum tolerances are $dL^* \leq 2$, $dC^* \leq 2$ and $dE_{(CIE 1976)} \leq 2$. When using the X-Rite reference data that was created with the 2014 version of the ColorChecker SG chart, these tolerances cannot even be met when using a spectrophotometer. When using a lighting set-up and a camera system, the color deviations will become even larger.

The implications of this work for attempts to meeting the FADGI guidelines [20] in combination with using X-Rite's reference data may be less dramatic, since FADGI does not specify tolerances for individual patches. Only an average tolerance is mentioned on the total set of 140 patches of the ColorChecker SG chart. This tolerance of $\text{CIEDE2000} \leq 2$ can be met with a camera profiling system using the reference data, although also in this case part of the tolerances have already been consumed by the deviations between ColorChecker SG charts.

Our results underline the necessity to use own spectrophotometric data for specific charts, instead of generic reference data from the supplier, when the highest color accuracy is aimed for. This is in line with the recommendations by the supplier of the ColorChecker SG chart, who states that “for the highest color accuracy” it is recommended to create individually measured custom reference data [47]. The analysis in this article not only supports that recommendation, but also quantifies it.

For example, we found that when using the publicly available X-Rite reference data in combination with an X-Rite ColorChecker SG chart produced after 2014–2015, the most strict Metamorfoze guidelines can only be met if the camera profiling process is manually tweaked. In this tweaking process typically the photographer manually changes parameters in the camera profile, after they were optimized by the software. This manual process is often very time-consuming and characterized by a succession of ad hoc decisions that makes it hardly reproducible for colleagues. The aim of this tweaking is to get a closer alignment between the generic reference data and the specific data produced by the camera profile for the individual X-Rite ColorChecker SG chart under the given lighting conditions. But since we have seen that the public reference data does not accurately describe X-Rite ColorChecker SG charts produced after 2014–2015, this manual tweaking process will not improve the actual color accuracy if the software is based on the public reference data. It is incorrect trying to optimize camera profiles by letting them reproduce reference data that are based on different black color patches than the black color patches of X-Rite ColorChecker SG charts produced after 2014–2015.

Our data show that when using the same reference data and X-Rite ColorChecker SG chart, satisfying the most strict FADGI guidelines may still be possible, although also in that case there is hardly any room left for deviations in lighting and camera set-up. Also in this case manual tweaking are better avoided.

We have shown that if the publicly available reference data are used for camera profiling, the largest color deviations occur for the black patches of the X-Rite ColorChecker SG chart. We have shown that in the digitized images, the deepest blacks would show the largest misrepresentations even on a calibrated display. On the other hand, if custom reference data is used for X-Rite ColorChecker SG charts produced
after 2014–2015, the deepest blacks cannot be obtained because they are not included in the chart. In that case, subtle nuances in dark colors on the original art object may get lost in digital reproductions. This would be unfortunate in for example the photography of 17th century Dutch paintings, where dark nuances form an important aspect of the art works.

Therefore it is indeed recommended to create individually measured custom reference data when using an X-Rite ColorChecker SG chart produced after 2014–2015. Our data show that for these measurements it is not necessary to use high-end spectrophotometers with prices usually exceeding €10.000. Instead, instruments such as the i1Pro at a 5 times lower price range were shown to qualify. We advise photography studios that utilize more than one X-Rite ColorChecker SG chart to create separate custom reference data files for each chart. This is also important when outsourcing parts of the digital photography to external photographers. Our recommendation implies that for building camera (ICC) profiles, it is preferred to use software such as basIColor [48] and Phocus [49] that allow the user to upload custom reference data.

Since we have shown that X-Rite ColorChecker SG charts produced after 2014–2015 no longer contain the very darkest color patches, with L*=8–9, we will investigate best ways to develop and include additional color patches, and how to apply these insights in cross-polarization photography.

Apart from that, we will also continue this investigation by developing a best practice for cases where the glossiness of the art object differs from the glossiness of the reference color patches. The current practice of camera profiling tacitly assumes that art object and X-Rite ColorChecker SG chart have the same glossiness. The current work has clearly demonstrated that even a relatively small difference in glossiness of 7 Gloss Units, on a scale that runs from 0 (dead matte) to 100 (high gloss), leads to color deviations large enough to make the profile fail the tightest guidelines, especially for black colors.

Declarations

Availability of data and material
All data generated or analysed during this study are included in this published article.

Competing interests
The authors declare that they have no competing interests.

Funding
The research reported in this paper is performed without separate funding.

Authors' contributions
EK, TK, HvB, CvW: coordination, conceptualization, wrote the main manuscript texts and analysed the overall results. All authors read and approved the final manuscript.

Acknowledgments
We thank Arthur Schmeling and Liane May (X-Rite) for support with ColorChecker charts. We wish to thank Sharath Balan and Nisha Goyal (AkzoNobel Bangalore), Margot Vlot and Joyce van Weerenburg-
Gigengack (AkzoNobel Sassenheim) for designing, formulating and preparing the dozens of very dark paint samples used in this investigation.

Authors' information
5. eric.kirchner@akzonobel.com, AkzoNobel Paints and Coatings, Color Technology, Rijksstraatweg 31, 2171 AJ Sassenheim, The Netherlands
*Corresponding author
6. c.van.wijk@rijksmuseum.nl, Rijksmuseum, Image Department, Museumstraat 1, P.O. Box 74888, 1070 DN Amsterdam, The Netherlands
7. h.van.beek@rijksmuseum.nl, Rijksmuseum, PK-online, Museumstraat 1, P.O. Box 74888, 1070 DN Amsterdam, The Netherlands
8. tammo.koster@akzonobel.com, AkzoNobel Paints and Coatings, Color Technology, Rijksstraatweg 31, 2171 AJ Sassenheim, The Netherlands

References
[1] Heyman S. A museum at the forefront of digitization. New York Times, 13 May 2015. https://www.nytimes.com/2015/05/14/arts/international/a-museum-at-the-forefront-of-digitization.html. Accessed 18 June 2020.

[2] Korytkowski P, Olejnik-Krugly A. Precise capture of colors in cultural heritage digitization. Color Res Appl. 2017;42:333–336.

[3] Saunders D. High-quality Imaging at the National Gallery: Origins, Implementation and Applications. Computers and the Humanities 1998;31:153–167.

[4] Noble P, Duijn E van, Hermens E, Keune K, Loon A van, Smelt S, Tauber G, Erdmann R. An exceptional commission: conservation history, treatment and painting technique of Rembrandt’s Marten and Oopjen 1634. Rijksmuseum Bull. 2018;66:309-334.

[5] Beek H van, Wijk C van. Operation Night Watch – High resolution imaging of Rembrandt’s largest painting. AHFAP Association for Historical & Fine Art Photography. Presentation for the 2019 AHFAP Conference, London.

[6] Smelt S, Erdmann R, Klein Gotink R, Noble P, Tauber G, van Wijk C. High-resolution imaging as a new research tool in the Rijksmuseum. Book of abstracts of the AIC 46 Annual meeting – Imaging Technology in Houston. Washington: American Institute for Conservation of Historic & Artistic Works:2018.

[7] Berns RS, Smith S. Analysis of color management default camera profiles for museum imaging applications. IS&T Archiving Conference, 2012: 1-5.
[8] Olejnik-Krugly A, Korytkowski P. Precise color capture using custom color targets. Color Res Appl. 2020;45:40-48. https://doi.org/10.1002/col.22435.

[9] Trumpy G. Digital reproduction of small gamut objects: a profiling procedure based on custom color targets. Proceedings of the Conference on Colour in Graphics, Imaging, and Vision (CGIV), 2010:143-147.

[10] Berns RS. The science of digitizing paintings for color-accurate image archives: a review. J Imag Sci Techn. 2001;45:305-325.

[11] FADGI - US Federal Agencies Digitization Guidelines Initiative, Still Image Working Group. Technical Guidelines for digitizing cultural heritage materials. The Library of Congress, Washington, 2010.

[12] Dormolen H van. Metamorfoze preservation imaging guidelines. National Library of the Netherlands; 2012. Version 1.0.

[13] Williams D, Burns PD. A decade of experience with digital imaging performance guidelines: the good, the bad, the missing. Proceedings of the Archiving 2017 conference, 165-169.

[14] Liu M, Konya I, Nandzik J, Flores-Herr N, Eckeler S, Ndjiki-Nya P. A new quality assessment and improvement system for print media. EURASIP J Adv Sign Proc. 2012, 109. https://doi.org/10.1186/1687-6180-2012-109.

[15] Martínez B, Mitjà C, Escofet J. Optimization of digitization procedures in cultural heritage preservation. Martins Costa MFPC (editor), Proceedings of SPIE, The International Society for Optical Engineering. 2013;8785:8785BD.

[16] Bruce A, Shifting standards. Presentation at Conference 2+3D Photography practice and prophecies, 9-10 May 2019, Rijksmuseum Amsterdam. https://www.youtube.com/watch?v=cCL_NOGM7_k. Accessed 19 June 2020.

[17] DTCH - Digital Transitions Cultural Heritage. Color Reproduction Guide. 2017.

[18] Zhao Y, Berns RS. Image-based spectral reflectance reconstruction using the matrix R method. Color Res Appl. 2007;32:343–351.

[19] Verhoeven G. Basics of photography for cultural heritage imaging. In: Stylianidis E, Remondino F, editors. 3D Recording, documentation and management of cultural heritage. Caithness: Whittles Publishing; 2016, p. 127-251.

[20] FADGI. Still Image Working Group. Technical Guidelines for Digitizing Cultural Heritage Materials, The Library of Congress, Washington. 2016.
[21] Golden Thread Object-level targets, available from Image Science Associates, http://www.imagescienceassociates.com. Accessed 19 June 2020.

[22] Golden Thread / DICE device-level targets, available from Image Science Associates, http://www.imagescienceassociates.com. Accessed 29 June 2020.

[23] Wyble DR. Next generation camera calibration target for archiving. Proceedings of the Archiving Conf. 2017:127-132.

[24] Wyble DR. Spectral implications for camera calibration target. Proceedings of the Archiving Conf. 2018:42-46.

[25] Avian Rochester and DT Cultural Heritage, The Next Generation Target. http://www.avianrochester.com/nextgentarget.php and https://dtculturalheritage.com/the-dt-generation-target-version-2. Accessed 30 June 2020.

[26] Mohammadi M, Nezamabadi M, Berns RS, Taplin LA. A prototype calibration target for spectral imaging. Proceedings of the AIC Colour 05 congress, 2005, 387-390.

[27] X-Rite, Color Checker Classic, available since 1976. https://xritephoto.com/camerasolutions. Accessed 26 June 2020.

[28] McCamy CS, Marcus H, Davidson JG. A color-rendition chart. J Appl Photogr Eng. 1976;2:95-99.

[29] Vitorino T, Casini A, Cucci C, Gebejesje A, Hiltunen J, Hauta-Kasari M, Picollo M, Stefani L. Accuracy in colour reproduction: using a Color-Checker chart to assess the usefulness and comparability of data acquired with two hyperspectral systems. In: Trémeau A, Schettini R, Tominaga S, editors. Proceeding of 5th international workshop on computational color imaging (CCIW). Heidelberg: Springer; 2015. p. 225–35. https://doi.org/10.1007/978-3-319-15979-9_21.

[30] Rodney A. Color management for photographers: hands on techniques for Photoshop users. Amsterdam: Elsevier; 2005.

[31] Yu DKC, Oulton DP. Colour difference acceptability for calibrated digital images. Imag Sci J. 2000;48:165-176.

[32] Abed FM, Berns RS. Geometry-independent target-based camera colorimetric characterization. J Imag Sci Tech. 2013;57:050503.

[33] Fraser B, Murphy C, Bunting F. Real world color management. 2nd ed. Berkeley: Peachpit press; 2005.

[34] X-Rite, Color Checker Digital SG, 2020. https://xritephoto.com/camerasolutions. Accessed 26 June 2020.
[35] Sanderson K. A comparative study of handheld reflectance spectrophotometers. Proceedings of
the AIC Annual Meeting Topics in Photographic Preservation Volume 16, AIC, 2015, 47-62.

[36] X-Rite, reference data for ColorChecker SG before and after November 2014.
https://www.xrite.com/service-
support/new_color_specifications_for_colorchecker_sg_and_classic_charts. Accessed 1 July 2020.

[37] Munsell Linear Gray Scale, https://munsell.com/color-products/color-standards/munsell-neutral-
value-gray-scales. Accessed 1 July 2020.

[38] Image Engineering, iQ-Analyzer software. https://www.image-
engineering.de/products/software/376-iq-analyzer. Accessed 24 September 2020.

[39] Beek H van, Wijk C van, Operation Night Watch – high resolution imaging of Rembrandt’s largest
painting. Conference of the Association for Historical & Fine Art Photography (AHFAP).
https://ahfap.org.uk/conferences#f8bc33eb-ad8b-4466-98e7-f6567c730edd. Accessed 1 July 2020.

[40] OpenDICE, http://www.digitizationguidelines.gov/guidelines/OpenDICE/Sample_Profile_ColorCheckerSG.xlsx.
Accessed 3 July 2020.

[41] X-Rite. Technical specifications for i1Pro. https://xritephoto.com/ph_product_overview.aspx?
ID=1913&Action=Support&SupportID=5586. Accessed 1 July 2020.

[42] X-Rite, Technical specifications for i1Pro2. https://www.xrite.com/categories/calibration-
profiling/i1basic-pro-2. Accessed 1 July 2020.

[43] Mouw T. Sphere vs. 0°/45°, a discussion of instrument geometries and their areas of application.
X-Rite 1995.
https://www.xrite.com/-/media/xrite/files/apps_engineering_techdocuments/s/sphere_vs_045_en.pdf.
Accessed 17 June 2020.

[44] Berns RS. Billmeyer and Saltzman's Principles of Color Technology. 4th ed. Hoboken; Wiley: 2019.

[45] BYK-Gardner, technical specifications of the spectro2guide instrument. https://www.byk-
instruments.com/es/en/Color/spectro2guide-Color-Control-Handheld-Spectrophotometer/Color-and-
lightfastness-with-digital-standards/spectro2guide%2C-45-0/p/7075. Accessed 9 July 2020.

[46] Natural Color System Gloss Scale, https://ncscolour.com/product/ncs-gloss-scale/. Accessed 6
July 2020.

[47] X-Rite. How to create custom reference data for ColorChecker Charts using i1Profiler?
https://www.xrite.com/service-
support/how_to_create_custom_reference_data_for_colorchecker_charts_using_i1profiler. Accessed 1 July 2020.

[48] BasICColor, Available at www.basiccolor.de. Accessed 19 June 2020.

[49] Phocus. Available at https://www.hasselblad.com/phocus. Accessed 19 June 2020.

Figures

Figure 1

X-Rite ColorChecker SG chart.

Figure 2

Munsell Linear Gray Scale.
Figure 3

Color deviations CIEDE2000 on the X-Rite ColorChecker SG chart between custom data and supplier’s reference data [36]. Colors indicate the magnitude of color error, with CIEDE2000 < 1 marked green, values between 1 and 2 marked light green, values between 2 and 3 marked yellow and values above 3 marked red.
**Figure 4**

(a) Comparison between measurements of $L^*$ by two 45°/0° instruments: i1Pro and spectro2guide. The dashed line corresponds to perfect alignment of the two instruments. (b) Difference in $L^*$ values as measured by the two instruments, as function of measured lightness value.
Figure 5

Difference in L* values as measured by the i1Pro and spectro2guide instruments, as function of measured gloss value at 60°.