https://doi.org/10.2352/lim.2024.5.1.24

Improving accuracy of color reproduction on mobile displays

Eric Kirchner[1], Lan Njo[1], Esther Perales[2], Aurora Larrosa Navarro[2], Carmen Vázquez[2], Ivo van der Lans[1] and Peter Spiers[1]
1. Department Surface Innovation & Analytics, AkzoNobel Paints and Coatings, Sassenheim, the Netherlands.
2. Department of Optics, University of Alicante, Alicante, Spain

Abstract

When visualizing colors on websites or in apps, color calibration is not feasible for consumer smartphones and tablets. The vast majority of consumers do not have the time, equipment or expertise to conduct color calibration. For such situations we recently developed the MDCIM (Mobile Display Characterization and Illumination Model) model. Using optics-based image processing it aims at improving digital color representation as assessed by human observers. It takes into account display-specific parameters and local lighting conditions.

In previous publications we determined model parameters for four mobile displays: an OLED display in Samsung Galaxy S4, and 3 LCD displays: iPad Air 2 and the iPad models from 2017 and 2018. Here, we investigate the performance of another OLED display, the iPhone XS Max. Using a psychophysical experiment, we show that colors generated by the MDCIM method are visually perceived as a much better color match with physical samples than when the default method is used, which is based on sRGB space and the color management system implemented by the smartphone manufacturer. The percentage of reasonable to good color matches improves from 3.1% to 85.9% by using MDCIM method, while the percentage of incorrect color matches drops from 83.8% to 3.6%.

Introduction

Over the past years the need has grown for photorealistic rendering. Online shopping has become a major sales channel. To improve customer satisfaction and to reduce the percentage of returns there is a growing need for accurate digital representation of objects.

Manufacturers of mobile displays use color management systems on top of the common device-independent sRGB color space [1] that best suits the users in general. These color management systems are not accessible to the user. Some smartphones and tablets have options for the user to change the color representations. For example, on Apple devices the user may choose to switch between light and dark appearance mode, or to switch on/off night shift and true tone. On many Samsung devices there are switches to choose either natural or vivid screen mode, light or dark mode, adaptive brightness, etc. But for all current mobile displays the user has only very limited options to change details of the color management system.

However, the requirements on color management for general use may be different from the requirements in specific uses. Manufacturers of mobile display individually balance the settings to for example optimize the readability of websites, the appearance of social media and to maximize user preferences for watching movie content. But the same settings are unlikely to also optimize the color reproduction accuracy when digital content is viewed in comparison to the corresponding object in the physical world. For example, when viewing a football match on the mobile display the user may prefer the grass to be greener than the actual grass in the stadium. But when ordering sneakers through an online website, that same user is more likely to return the product after delivery if its physical world color does not match the color expected from the online experience.

There are several commercial solutions available for color calibrating computer displays. These are mainly used by professional users because of the costs, expertise and time needed for calibration. For mobile displays such options are hardly available, and certainly not for general consumers. Therefore there is a need for a fast and easy method that does not need specialized equipment or expertise, but that does improve the color reproduction accuracy when visually comparing digital color representations with the corresponding objects in the physical world.

For this goal we recently developed the Mobile Display Characterization and Illumination Model (MDCIM) model [2][3]. Based on an analysis of optics it is implemented as image processing. Experimental validations show that it achieves higher color accuracy in cross-media color reproduction on several displays. This was confirmed on three representative LCD displays (iPad Air 2, iPad models from 2017 and 2018) [2] and one OLED display (Samsung Galaxy S4) [3]. These results were obtained by accounting for two factors that contribute most to color deviations in cross-media color reproduction.

(1) Display-specific color deviations. Depending on the colorimetric specifications of the display and the parameters of the color management system used in the operating system of the mobile device, the spectrum of the light emitted by the display may be more or less similar to the spectrum of the light that is reflected from the physical object it intends to reproduce.
(2) Variations in ambient light. Even if the ambient light that is reflected from the physical object has a spectrum equal to the illuminant (usually standardized daylight D65) that is assumed in calculating the color representation, for creating an accurate color representation the illuminance ("intensity") of the ambient light is a crucial factor that is not accounted for yet.

The physics-based derivation of the MDCIM model can be found in previous publications [2][3]. In the current conference proceeding, we test the accuracy of the MDCIM model for a different (newer) OLED display, the display of the iPhone XS Max. In the experimental section below we discuss the relevant MDCIM model parameters and the spectroradiometer measurements needed to derive their optimum values. We also describe the visual experiment that we used to validate the MDCIM model and its parameter values for the display of the iPhone XS Max. The next section discusses the results from this psychophysical test. In the final section we summarize our conclusions and recommendations for further work.

Experimental

MDCIM model parameters

In the MDCIM model the relation between tristimulus values X, Y, Z (which are directly related to CIELab colorimetric parameters) and the luminance values of the red, green and blue color channels Y_R, Y_G and Y_B is given by a 3x3 matrix M, not unlike the GOG-model from Berns [4][5]. These luminance values are assumed to behave exponentially. For example, for the red channel in eight-bit representation a value R ranging from 0 to 255 results in a digital value $d_R = R/255$ and a luminance of the red channel of

$$Y_R = \left(k_{1,R} d_R + k_{2,R}\right)^{\gamma_R} \quad (1)$$

At low luminance values $d < d_0$ we assume the luminance to scale linearly with a coefficient β. Once the best fit values for parameters d_0 and γ are determined, the values of k_1, k_2 and β directly follow as a result of mathematically enforcing continuity of the luminance function and its derivative [2][3]. The tristimulus values just mentioned are all relative to the corresponding values X_0, Y_0, Z_0 of a completely black image, which is important especially when characterizing LCD displays. The display of the iPhone XS Max that is investigated here is based on OLED technology, for which the values X_0, Y_0, Z_0 are expected to be very small.

In the MDCIM model matrix M then assumes the following form (for details and mathematical derivation see [2][3], where also the method to take into account ambient lighting is discussed):

$$M = \begin{pmatrix} \frac{x_R}{y_R} Y_{R,\max} & \frac{x_G}{y_G} Y_{G,\max} & \frac{x_B}{y_B} Y_{B,\max} \\ Y_{R,\max} & Y_{G,\max} & Y_{B,\max} \\ \frac{z_R}{y_R} Y_{R,\max} & \frac{z_G}{y_G} Y_{G,\max} & \frac{z_B}{y_B} Y_{B,\max} \end{pmatrix} \quad (2)$$

This shows that the matrix elements are related to the chromaticity coordinates x, y and maximum luminance values Y_{max} of the red, green and blue channel. For sRGB space, M is given by:

$$M_{sRGB} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix} \quad (3)$$

This enables us to derive the values of chromaticity coordinates and maximum luminance tacitly assumed in sRGB color space.

Spectroradiometer measurements

For optimizing the display specific model parameters, we followed the same process that we used for the previous displays [2][3]. In short, we created 55 images with various R,G,B values. For example, to determine the parameters of the red color channel we created images with G=B=0, whereas R=0, 5, 10, …, 45, 50, 75, 100, 125, …, 225, 255.

Table 1: Spectroradiometer data for iPhone XS Max display, and corresponding values implicit in sRGB space.

	sRGB	iPhone XS Max
x Red	0.64	0.6457
y Red	0.33	0.3321
Y_{max} Red	21.26	118.66
x Green	0.30	0.3101
y Green	0.60	0.5938
Y_{max} Green	71.52	373.27
x Blue	0.15	0.146
y Blue	0.06	0.0609
Y_{max} Blue	7.22	37.61
x Black	0.0	0.2941
y Black	0.0	0.2639
Y Black	0.0	0.0

In a completely dark room, we measured the spectrum of the light emitted by the display when showing each of the 55 images, resulting in chromaticity coordinates x, y and luminance values Y. This was done using a CS-2000A spectroradiometer (supplier Konica-Minolta) with its aperture mounted against the display. The display was set at maximum brightness, and automatic adjustment of display brightness was switched off.

By fitting the spectroradiometer measurement data to equations (1), (2) and (3) we obtained optimized values for the MDCIM model specifically for the iPhone XS Max display [6]. The results are shown in Table 1 and 2.

Table 2: MDCIM model parameters based on spectroradiometer data for iPhone XS Max. These are compared to corresponding values assumed in sRGB space.

	sRGB	iPhone XS Max
Red channel		
γ	2.4	2.231
d_0	0.04045	0.00784
k_1	0.947867	0.990441
k_2	0.055	0.009559
β	0.077399	0.015000
Green channel		
γ	2.4	2.255
d_0	0.04045	0.00784
k_1	0.947867	0.99026
k_2	0.055	0.00974
β	0.077399	0.01394
Blue channel		
γ	2.4	2.287
d_0	0.04045	0.00784
k_1	0.947867	0.99001
k_2	0.055	0.00999
β	0.077399	0.01264

Psychophysical experiments

We conducted visual tests with 8 observers (3 men, 5 women), all tested to have normal color vision [6]. In order to estimate repeatability, each observer conducted the test three times on different days.

In each test we used 30 physical samples from the NCS (Natural Color System) Index 1950 (Quality level 2) [7]. In order to cover color space, we selected 4 different colors of various saturation and from each of the 6 main color categories (red, orange, yellow, green, blue and purple), as well as 6 achromatic colors with varying lightness (Table 3). All samples were matte, in order to make the color evaluation less dependent on viewing angle and lighting.

For each sample, reflectance with 10 nm resolution was measured with a spectrophotometer (BYK-mac i, supplier BYK Gardner). We used data from the 110º measurement geometry in order to minimize the influence from specular gloss. Using standard colorimetry [8], reflectance data was converted into tristimulus values X, Y, Z assuming D65 standardized daylight illuminant and standard 2 º observer. With the equations (1), (2) and (3) these tristimulus values were converted into either default RGB values (using sRGB color space) or RGB values based on the MDCIM model. As a result, we obtained two RGB images for each of the 30 physical samples.

The psychophysical test was executed in a light booth, the JUST Normlight LED Color Viewing Light. This light booth is used in several industries related to textiles, plastics and automotive. The light booth enables us to choose certified D65 lighting in our experiment. We chose an illuminance level of 750 lux to simulate typical viewing conditions for indoor applications.

The experiment is illustrated in Figure 1. During the experiment we asked observers to hold a physical sample next to two images that were first simultaneously shown on the display (Figure 2, left). One of the images was generated using the default method (involving sRGB color space and the default color management system on the mobile device), whereas the other image was generated using the MDCIM method as a post-calculation. Both calculation methods were randomly distributed over the top and the bottom images, so observers did not know which method generated which image. The color surrounding these images was made such as to best simulate the color of the floor part of the light booth, in order to reduce effects from simultaneous contrast.

Observers assessed the accuracy of the visual color match between a displayed color and a physical sample. For this, they used a scoring table as shown in Table 4. We asked the observers the three following questions:

(1) Which of the two images shows the best color match as visually compared to the physical sample? (Figure 2, left)
(2) What is the visual score for the color match between the top image and the physical sample? (Figure 2, center)
(3) What is the visual score for the color match between the bottom image and the physical sample? (Figure 2, right).

Figure 1. *Psychophysical experiment: general set-up.*

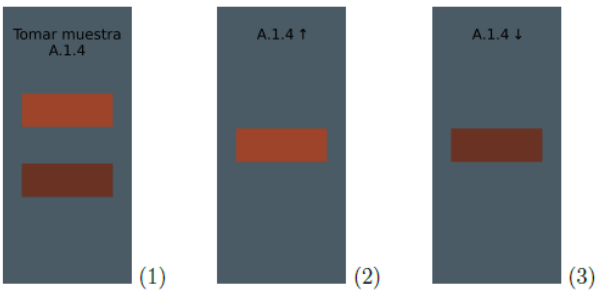

Figure 2. *Three consecutive screens during the psychophysical experiment.*

Table 3: Physical samples used for psychophysical experiments.

NCS Code	L*	a*	b*	NCS Code	L*	a*	b*
S1005-R10B	85.4	2.8	1.9	S6030-B90G	30	-23.8	3.2
S5030-R10B	36.7	22.6	3.9	S1020-B10G	79	-11.9	-7.2
S0540-Y80R	73.7	29.1	19.2	S1050-B30G	69.3	-32.4	-12.6
S3060-Y70R	40.7	36	33.6	S1040-R70B	70.4	1.7	-23.8
S0520-Y30R	85.2	8.6	23.7	S2020-R90B	70.6	-5.7	-11.8
S0560-Y40R	72	30.5	48.8	S2010-R30B	73.3	5.7	-2.2
S1070-Y20R	69.7	22.9	67.6	S2030-R50B	67.2	12.6	-13.9
S3030-Y30R	60.5	12.8	28.5	S4050-R60B	28.7	19.4	-36.4
S0570-Y10R	78.6	15.7	75.7	S5020-R50B	44.6	9	-12
S3040-Y00R	60.6	4.2	41.9	S2000-N	77.2	-0.1	1.2
S0520-G90Y	88.4	-2.4	26.5	S7500-N	34.5	0.1	-1.2
S4040-G90Y	51.3	2.1	38.3	S3500-N	66.6	0	0.1
S3060-G30Y	46.2	-23.7	41.2	S5500-N	50.5	-0.1	-0.6
S4020-G10Y	55.4	-13.4	7.7	S1502-Y	81.6	-0.5	4.9
S3040-B80G	55.6	-29.5	0.8	S4052-B	57.6	-1	-2.4

Table 4: Scoring table for visual assessments during the psychophysical test.

Visual score	Description
0	No/Hardly any difference
1	Small, negligible difference
2	Difference visible but still reasonable
3	Difference visible, doubtful match
4	Difference clearly visible, not correct match
5	Large difference, very bad match

Results

Visual test results

Repeatability and reproducibility

Since 8 observers conducted the test on 30 physical samples, and conducting each test three times independently, we obtained a total of 720 visual evaluations for the images created by the MDCIM method, and 720 visual evaluations for the images created by using the default method which, as mentioned before, is based on sRGB color space and the color management system implemented by the smartphone manufacturer.

We first determined the intra-observer repeatability. Since every observer did the same test three times independently of each other, and with the order of images randomized for each session, we collected three independent visual scores (according to Table 4) for each color and each observer and evaluating color accuracy for both, sRGB and MDCIM. This allows us not only to determine the average visual score for each observer-color combination, but also the standard deviation in the visual scores for each observer-color combination. This standard deviation is a measure for the intra-observer repeatability of the psychophysical test.

Table 5 shows that for the default method, intra-observer repeatability is 0.40. This means that 68% of the visual scores given with the descriptions of Table 4 are exactly repeated by the same observer in an independent session. When using the images created by the MDCIM method the repeatability slightly increases to 0.47.

Intra-observer repeatability refers to the consistency in visual scores for a single observer as compared to that same observer. We also determined the inter-observer reproducibility. This is the consistency in visual scores for a single observer as compared to each of the other observers. It is calculated for each observer-color pair by calculating the absolute difference between the average of the three sessions of that single observer and the average of the three sessions of all eight observers. Table 5 shows that inter-observer reproducibility of the default method is 0.41, while for the MDCIM method it is 0.58. This shows that as expected the consistency in results between different observers is higher than the consistency in results from an observer with the same observer.

Since all visual scores in these tests are given on a five-point scale (Table 4), the current results can be considered to represent good repeatability and reproducibility. These results are important as a reference when further analyzing the results of the psychophysical experiments. Deviations in visual scores larger than the repeatability and reproducibility should be given full consideration.

Table 5: Intra-observer repeatability and inter-observer reproducibility for iPhone XS Max.

Parameter	Default	MDCIM
Intra-observer repeatability	0.40	0.47
Inter-observer reproducibility	0.41	0.58

Preference for visualization

The first question asked to the observers was which of the two images, generated in randomized order with the default method and with the MDCIM method, had the best color match when compared with the physical sample.

Our results show that in 98% of all the assessments the image generated with the MDCIM method was preferred.

Perceived accuracy of digital color representation

From the answers to the second and third question asked during the psychophysical experiment, we calculated the average visual score for the images created by each of the two methods. When using the default method based on sRGB color space and the color management system provided by the smartphone manufacturer, the average visual score is 4.3. According to Table 4, this can be interpreted as observers finding the average color match between digital color representation and the physical sample not satisfactory. The average color match is perceived as not being a correct color match.

When using the MDCIM method instead for generating digital color representations, the average visual score reduces to 1.3. This can be interpreted as a small, negligible color difference and certainly reasonable color match. Our results show that not only do observers strongly prefer the visualization by the MDCIM method over the default method, they are also on average finding the resulting color match with the physical sample satisfactory when using the MDCIM method and unsatisfactory when using the default method. The improvement in visual score is much larger than the intra-observer repeatability and the intra-observer reproducibility.

Table 6 shows that the percentage of visual scores that are 4 or larger drops from 83.8% to 3.6% by using the MDCIM method. In other words, the percentage of incorrect color matches strongly improves by using the MDCIM method. Similarly, the percentage of reasonable to good color matches strongly improves by using MDCIM method. This percentage increases from 3.1% to 85.9%.

Table 6: Visual scores for default images and images generated by the MDCIM method on iPhone XS Max.

Parameter	Default	MDCIM
Average visual score	4.3	1.3
Percentage with average visual score ≥ 4	83.8%	3.6%
Percentage with average visual score ≤ 2	3.1%	85.9%

Conclusions

In this conference proceeding we show the results of using a previously derived method, the MDCIM method, on an iPhone XS Max display. It requires the use of specialized equipment (spectroradiometer) only by developers of websites and apps, but not by the end-user. Using a psychophysical experiment, we show that colors generated by this method are visually perceived as a much better color match with physical samples than when the default method is used, based on sRGB space and the color management system implemented by the smartphone manufacturer. In 98% of the assessments the color representation calculated with the MDCIM method was preferred. We expect the improvement in color accuracy presented here to is also valid when digitally presenting complex scenes and color mixtures.

References

[1] International Electrotechnical Commission (IEC), 100/PT61966(PL) 34, International Standard, “Multimedia systems and equipment–colour measurement and management, part 2.1: default RGB colour space–sRGB” (1999), Amendment 1 (2003).

[2] E. Kirchner, I. van der Lans, F. Martínez-Verdú, E. Perales. "Improving color reproduction accuracy of a mobile liquid crystal display". J Opt Soc Am. A, 34, 101-110 (2017).

[3] E. Kirchner, I. van der Lans, E. Perales, F. Martínez-Verdú. "Improving color reproduction accuracy of an OLED-based mobile display". Color Res Appl., 43, 34–46 (2018). https://doi.org/10.1002/col.22148.

[4] R.S. Berns, R.J. Motta, M.E. Gorzynski, “CRT colorimetry, Part I: Theory and practice.” Color Res. Appl., 18, 299–314 (1993).

[5] R.S. Berns, “Methods for characterizing CRT displays.” Displays, 16, 173-182 (1996).

[6] A. Larrosa, Caracterización colorimétrica de dispositivos de visualización de datos. (MSc Thesis, June 2023, University of Alicante Spain).

[7] NCS, https://www.ncscolour.no/images/pdf/Spesifikasjon-NCS-INDEX-1950-Original.pdf

[8] http://www.brucelindbloom.com/