

Evaluating Color Transformation Quality and Accuracy of Prosumer and Mobile Phone Cameras for High Dynamic Range Cultural Heritage Documentation

Leah Humenuck and Mekides Assefa Abebe, Rochester Institute of Technology; Rochester, New York/USA

Abstract

The capability for a camera to produce a color-accurate high dynamic range image is based upon its capture of luminance within a scene as well as the targets chosen to create the color transformation matrix. The wide-ranging luminance within a scene is an important part of cultural heritage documentation to appropriately capture an object's appearance. In addition, color accuracy is critical to documenting cultural heritage appropriately. This research compares prosumer and mobile phone cameras for cultural heritage documentation using single exposure and high dynamic range images. It focuses on the evaluation of color characterization process, color reproduction quality, and generation of the scene. This was done by using two types of prosumer cameras and two types of mobile phone cameras, at 800 lux with a wide range of color targets with various surface textures, matte, semigloss, and glossy. It was found that for creating color calibration matrices, a single exposure outperformed the one created from fusion of multiple exposure images. Additionally, including an extended achromatic scale along with the traditional Macbeth colors as part of the training data for the color calibration matrix may increase color accuracy for different cameras and the samples in the scene.

Introduction

High dynamic range imaging (HDR) is a field which has made large advancements and is now extending application to cultural heritage. This type of imaging mimics human vision in perception of extremes of very bright and very dark regions of a scene at once [1]. Previously in the history of analog and earlier digital camera technology, the very bright regions or the very dark regions or neither of a scene could be captured properly. Possibly time consuming or added steps in post processing would need to be performed in order to show both of these extremes in the final image. However as digital cameras progressed, HDR imaging was developed and also became a setting where the camera would take a series of photos at various exposures and perform the necessary computations to output a scene which showed the detail at both the light and dark extremes. There are different ways in which to combine the series of images in order to create a final HDR image. These different ways of processing can produce different effects depending on what the optimal goal is for the HDR image. Regarding all of this though is the importance of understanding the possible range of luminance from that camera which are needed to produce an HDR image. The RAW format is a digital negative where there is no compression of data such as in other formats. The sensors in different types of cameras will be able to capture a different amount of luminance from each of these scenes and the quantized digital representation is stored in the RAW image file.

Adding to camera technology are mobile phone cameras, which historically only produced compressed file formats of their images until the early 2010s [2]. This has limited the amount of

research which could be performed on them as compressed file formats as well as proprietary computational photography obscured understanding of them. As mobile phone camera technology has progressed, demand for a manual setting increased and there are now either native or third-party apps which allow for custom settings and with these digital negative formats of their image files known as Digital Negative Image (DNG). This development of mobile phone cameras allows them to be compared in some aspects to prosumer cameras.

Currently for HDR imaging research, there is no international standard by which to judge HDR image capture. In lieu of this lack of standardization, cameras may still be assessed relative to one another by understanding their response to a set of representative targets and comparing the results to the ground truth measurements from a spectroradiometer.

This research investigates luminance ranges that prosumer cameras and mobile phone cameras capture for application to cultural heritage and the types of reflectance targets chosen to create the color calibration transformation matrix for color accuracy [3][4][5]. It is known that high dynamic range imaging differs from traditional imaging, by merging various exposures and increasing the range of luminance captured from a scene. The luminance values of a scene may extend beyond the lightest and the darkest patches found in the traditional Macbeth achromatic scale which is traditionally used in cultural heritage documentation. This challenge has been found in cultural heritage documentation where the details within dark regions may be lost using traditional imaging, and where color accuracy is of high importance [6]. Additionally, mobile phones have begun to be included in other industries and there is increased interest for them to be accepted as tools for documentation in cultural heritage. Because of the complexity of HDR imaging and the goal of accurate documentation, it is necessary to understand the capabilities of mobile phone cameras and evaluate techniques for HDR documentation. This research compares prosumer and mobile phone cameras within a reflectance scene at various exposures and compares the color transformations derived from a single exposure image and a multi-exposure fusion image (MEF) [7] as well as expanding the training set for the color calibration matrix and assessing it on other color targets.

Methodology

In this work, color reproduction capabilities of different camera technologies have been evaluated. Color reproduction on different reflective materials were assessed; matte, semi glossy, and glossy. It was important to have a variety of material surfaces in the scene in order to replicate likely scenes for art. The prepared scene, with a detailed diagram of the included targets, is shown in Figure 3. The captured color values of the reflective targets are compared with the spectroradiometer measurements. In order to convert the camera RGB space values to the device independent

XYZ space, several characterization approaches have also been evaluated. The characterization was performed using different combinations of subsample color targets from the single exposure and MEF images of the scene. A summary of the evaluated approaches is provided in Table 2. The employed experimental scene, measurement setup, color characterization approaches, as well as color difference comparison processes are presented as follows.

Scene Acquisition

Lighting

In order to assess the dynamic range of the two cameras, the targets were imaged in black painted room illuminated by LEDMotive Spectra Tune lights. The geometry of the lights and the camera were 45:0 as per the CIE recommended geometry [6]. The light sources were set at 100% intensity and have 800 lux illuminance and 5452K CCT, as measured with the MK350N Premium Spectroradiometer, see Table 1. The spectral power distribution was measured by a CS2000 spectroradiometer on a perfect reflecting diffuser within the scene, see Figure 1.

Table 1. Lux and CCT of Lights

Light Level Percentage of Intensity	100%
Lux	800
CCT	5452

Cameras

The prosumer cameras used were a Canon EOS 5D Mark II (C1) and a Sony ILCE-7RM3 (C2). The mobile phones used were a Huawei LYA-L29 (C3) and an iPhone 12 Pro (C4). Each camera was placed above the targets such that the targets filled the screen. Both prosumer cameras were operated remotely through a computer using proprietary software. C3 and C4 cameras were captured directly by adjusting the settings on the phones' screens with a finger.

The C3 images were captured using the native camera app and its corresponding settings, which includes a DNG file and manual exposure settings. The C4 images were captured using a third-party app downloaded from the Apple Applications Store, RAW+. This is due to the fact that the current native app does not allow for manual adjustment of the exposure settings, though it does take a DNG file. For both cameras the standard lens was used.

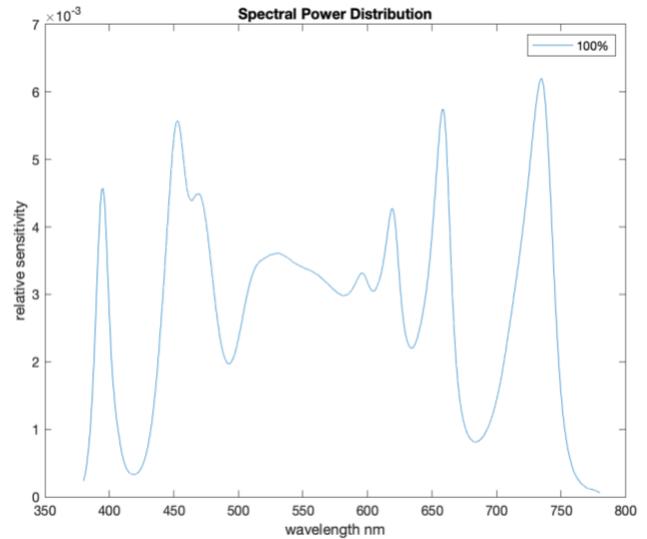


Figure 1. Spectral power distribution of lighting condition

Each camera took 32 images in RAW mode of the targets with a maximum exposure time of 1s and minimum of 1/4000s at 1/2s step size. This range of exposure times was chosen due to the limitations of the C3 and C4 cameras. Though each of the prosumer cameras can perform more exposures, to compare the capabilities of each camera to each other, it was needed to capture images within a shared exposure range. The C1 camera image size is 3752x5634 and 14-bit. The C2 camera image size is 5320x7986 and 14-bit. The C3 camera image size is 5472x7296, according to investigation the data appeared 12-bit. The C4 camera image size is 3024x4032 and 12-bit. The responses of the green channels of each of the cameras for the Munsell Glossy N6.5 patch are shown in Figure 2. The plot shows the limited dynamic range capabilities of the smartphone cameras compared to that of the prosumer cameras. The sensor responses of C3 and C4 are shown to be saturated at 1/5s of exposure, while the C1 and C2 show more potential with no saturation, even for longer than 1s exposure time.

Targets

The experimental scene is made up of targets, constituted a range of value, hue, and surface textures of matte, semigloss, and glossy. It included the Xrite ColorChecker SG (CCSG), Xrite ColorChecker Passport (PP), Next Generation Target (NGT), Munsell Linear Gray Scale (MLGS), Munsell Glossy achromatic samples (Achromatic or AC) every whole and half value from N0.5 to N9.5, a perfect reflecting diffuser, and five additional black samples of a spectrophotometer black trap, Musou black, Acktar spectral black, Acktar metal velvet, and telescope flock. A map of the targets as to how they were laid out for imaging is shown in Figure 3.

Image Processing

For each camera image, the raw files were converted into uncompressed 16-bit tiffs using RawDigger without scaling. The RGB values of a region of interest (ROI) on each patch were extracted and averaged. The colored mask of the identified ROIs is shown in Figure 5.

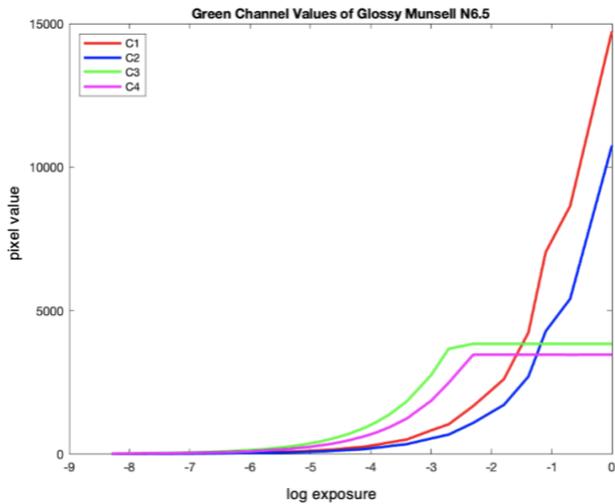


Figure 2. Green channel responses of each camera over all exposures for the Glossy Munsell N6.5.

For the single exposure, the RGB values were converted to 0-1 range by using the max pixel value. For the MEF images, the RGB values were converted to 0-1 range by using the max value from the diffuse white and any higher values were clipped to 1. This was performed such that pixel values higher than the diffuse white are considered as highlights in standard encoding format of an HDR images, and the analysis of highlights falls outside the scope of this paper. The measured XYZ values were normalized such that Y of the diffuse white was equal to 1 [8].

Each camera's RGB values were converted to XYZ by using a color calibration matrix, a 3x3 transformation matrix. Four different ways of camera characterization approaches, with a designated name in parenthesis, are investigated; 1) using the single exposure Macbeth colors from the Xrite ColorChecker SG (Traditional), 2) using the single exposure Macbeth colors from the Xrite ColorChecker SG plus the Achromatic samples (Modified #1), 3) using the Macbeth colors from the Xrite ColorChecker SG created from MEF (Modified #2), and 4) using the Macbeth colors from the Xrite ColorChecker SG plus the Achromatic samples created from MEF (Modified #3). A summary of the color calibration matrices' variables is found in Table 2. An illustration of the chosen patches is shown in Figure 5.

These targets and their ranges were chosen to represent a standard which is available to many cultural heritage imaging studios as well as to obtain samples which may provide darker values and higher which might be found in reflectance items of high dynamic range images. Additionally, we also chose to include targets which provided more intermediary achromatic steps than what is traditionally presented in commonly used color targets. The measured luminance values of all the achromatic patches are shown in Figure 4.

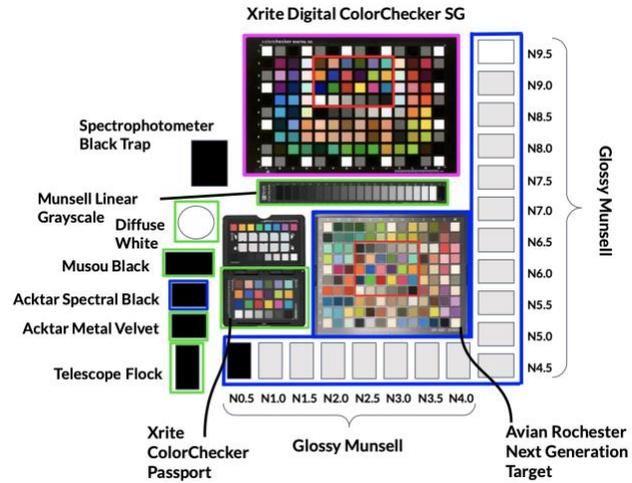


Figure 3. Schematic of targets and layout on the table. Targets highlighted in blue are glossy, highlighted in green are matte, and highlighted in magenta are semigloss. Not all the patches for Avian Rochester Next Generation Target and the Xrite Digital ColorChecker SG were used in this research, Only the patches highlighted in red were used.

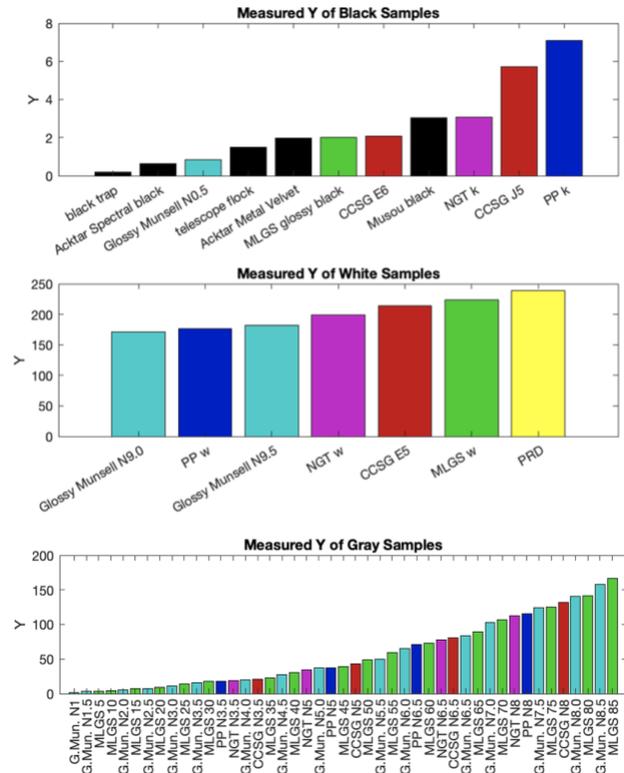


Figure 4. Measured luminance values of all achromatic samples. Top: all black-labeled samples. Middle: white-labeled samples. Bottom: all gray-labeled samples. They are color coded the same as the ROIs shown in Figure 5, with the added black samples, not found in commercial targets, in black.

Color Calibration and Comparison

For Traditional and Modified #1, the highest exposure image without overexposed pixels was chosen for each of the cameras: C1 - 1/4s; C2 - 1/3s; C3 - 1/40s; C4 - 1/40s. To generate the MEF image, Banterle et. al. 's implementation of Debevec's linear fusion method was utilized [9-10]. After subtracting the dark signal from each image at the same exposure, the process for deriving these matrices is found in through Equation 1, where n is the number of patches in the training set. The averaged RGB values of each target patch was converted to XYZ tristimulus values (XYZ_{estim}) using the resulted CCM and converted into CIELab values for further optimization.

$$CCM = (C^T C)^{-1} C^T D \quad (1)$$

Where, D and C represent the measured XYZ and camera RGB response values of the calibration target patches.

$$D = \begin{bmatrix} X_{P_1} & Y_{P_1} & Z_{P_1} \\ X_{P_2} & Y_{P_2} & Z_{P_2} \\ \vdots & \vdots & \vdots \\ X_{P_n} & Y_{P_n} & Z_{P_n} \end{bmatrix} \quad C = \begin{bmatrix} R_{P_1} & G_{P_1} & B_{P_1} \\ R_{P_2} & G_{P_2} & B_{P_2} \\ \vdots & \vdots & \vdots \\ R_{P_n} & G_{P_n} & B_{P_n} \end{bmatrix}$$

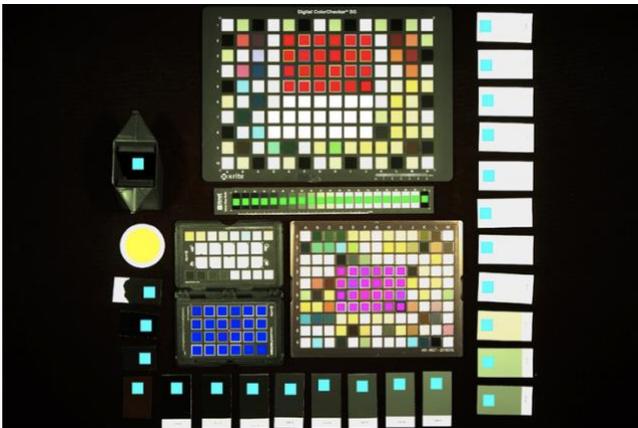


Figure 5. Example of the regions of interest (ROI) chosen on each of the targets. The ROIs are highlighted in red, blue, green, yellow, magenta, and cyan.

Table 2. Color Characterization Matrices

Name	Image	Training Data
Traditional	Single Exposure	Xrite ColorChecker SG Macbeth Colors (Red mask in Fig 5.)
Modified #1	Single Exposure	Xrite ColorChecker SG Macbeth Colors + Achromatic samples (Cyan and Yellow masks)
Modified #2	Multiple Exposure Fusion	Xrite ColorChecker SG Macbeth Colors
Modified #3	Multiple Exposure Fusion	Xrite ColorChecker SG Macbeth Colors + Achromatic samples

The measured tristimulus values (XYZ_{meas}) of the patches were collected with the CS2000 spectroradiometer, using the CIE 1931 standard observer. The CCM matrix is then optimized using

a least mean squares minimization of the DE2000 between the estimated XYZ_{estim} and the measured XYZ_{meas} of up to 8000 iterations.

Results

The predictions of each of the CCMs for all cameras was compared to the measured values of all target groups. The mean of all target groups for each camera under each CCM is found in figure 6. This shows how accurately the various CCMs were throughout each of the target groups. The Traditional and the Modified #1 were found to be the more accurate of the four CCMs. The mobile phones seemed to have benefited the most from the addition of the expanded achromatic range added to the training data as can be seen by their changing mean ΔE_{00} (Modified #1 and #3) for creating a CCM. This may be due to the limitations on their sensors and by supplying more training data for the CCM, it allows them to increase performance.

The ΔE_{00} for each of the Achromatic targets at the single exposure and the MEF are presented in Figure 7. This shows the performance of CCM generated from single exposure and MEF image targets. The mobile phones showed higher color differences for whiter patches compared to the prosumer cameras. The prosumer cameras remain more consistent than the mobile phone cameras in both the single exposure and the MEF images across all CCMs. The mobile phone cameras for the MEF images also varied more in the lower value achromatic patches N1.0-4.0. This effect is also shown in Figure 8 in the Glossy Black plot, where the both C3 and C4 icons are further from the measured value than the C1 and C2. Similarly, the CCMs from MEF generally performed worse. Such higher color differences in the white patches may have something to do with the over-exposure and clipping related issues.

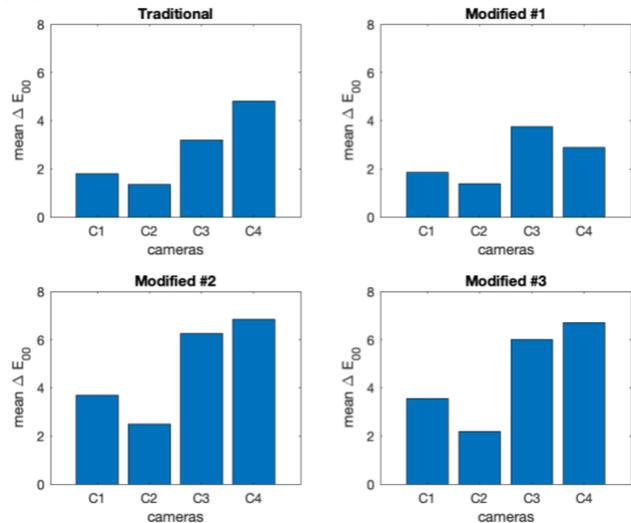


Figure 6. Mean ΔE_{00} across all targets for each camera with each CCM.

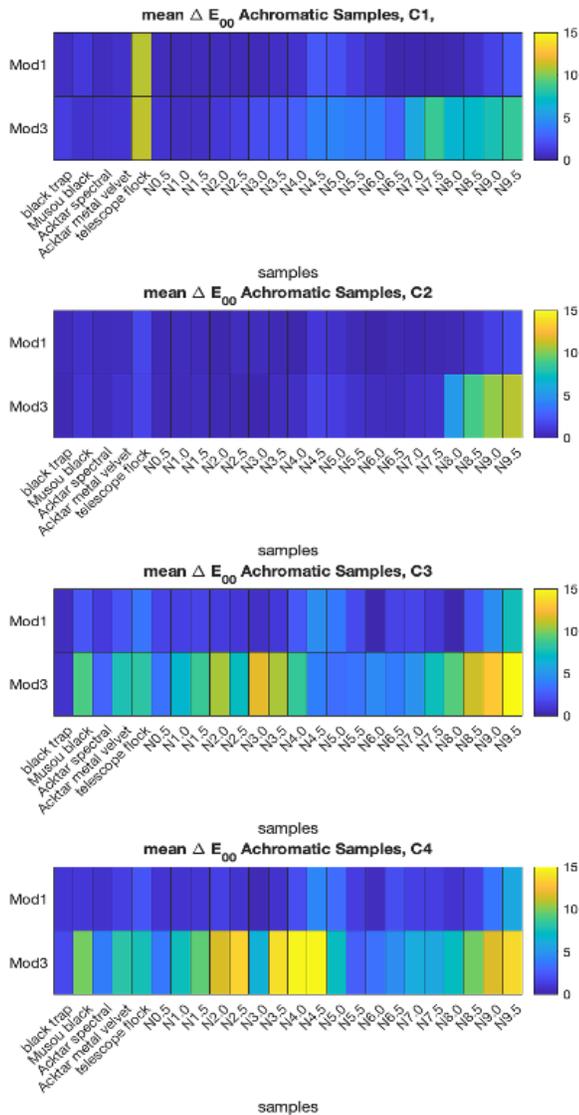


Figure 7. Mean ΔE_{00} of the Achromatic samples for Modified #1 and Modified #3 CCMs

Since the C4 was being operated by a third-party app, it may be that the controls of the camera are not to the same consistency and degree of accuracy as the native apps or controls on the other cameras. This could be a contributing factor as to why the results of C4 are not as consistent as the other cameras. Other third-party apps had been tested such as Manual Cam, MCCamera, Camera+, Manual, Pro Camera, and ProCamera among others. However, it was found that their interfaces were either difficult to work with for repeated use or did not provide the desired features such as a wide range of exposure values.

Figures 8 and 9 show the CIE Lab values of the cameras for all CCMs. In the black plots, the MEF images of the mobile phone cameras (Modified #2 and #3) showed that they were consistently at higher L values than the single exposure (Traditional and Modified #1) as well as the prosumer cameras at both MEF and single exposure. The MEF of all white and N5 were consistently varying in the red-green (a) or the yellow-blue (b) more than the L

values. The diffuse white based normalization and the clipping steps of processing may have contributed to the higher dark levels and the color shift of the MEF based mobile phone results. To fully understand all the factors contributing to the lower performances of the mobile phone cameras as well as the MEF results, further analysis is required.

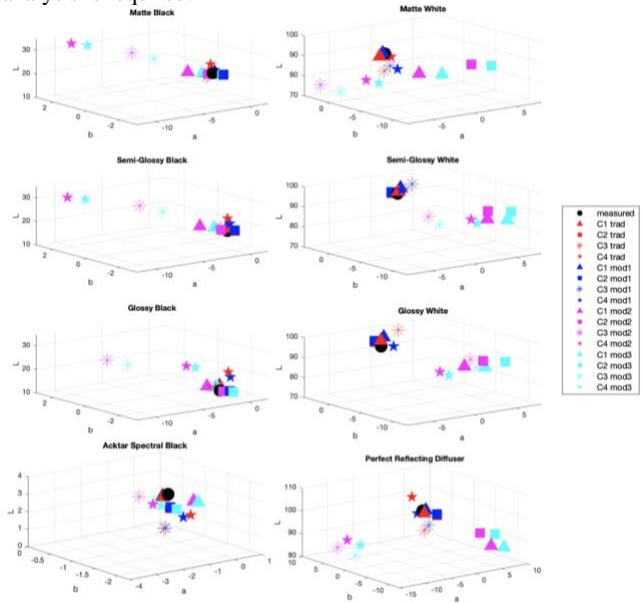


Figure 8. CIE Lab comparison of the black (right column) and white (left column) patches at various textures, measured versus the cameras at all CCMs.

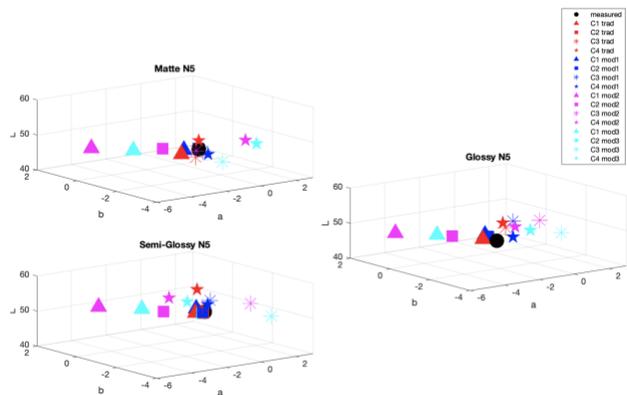


Figure 9. CIE Lab comparison of the N5 at various textures, measured versus the cameras at all CCMs.

Discussion

In general, our analysis showed the color reproduction limitations of mobile phone cameras compared to that of the prosumer cameras under extended ranges of sample types. Given the inherent differences in their sensor dynamic range capabilities, such performance is understandable. Both prosumer cameras also performed within the Federal Agencies Digital Guidelines Initiative Technical Guidelines for Digitizing Cultural Heritage Materials (FADGI) for their mean ΔE_{00} differences in both the single and MEF images cases [11].

It was interesting to note that the MEF caused all cameras to have lower color accuracy. A possible explanation as to why the MEF images had overall lower color accuracy (Figures 6 and 7) than the single exposure fusion images may be due to various factors, one of which may be a white point misalignment. During our analysis, we have observed areas with higher intensity values than the diffuse white. Therefore, the utilization of the diffuse white as a normalization factor might have caused over-exposures and hence more color differences for the MEF based predictions. The Debevec-Malik MEF method should also be investigated further and enhanced for better color accuracy. It is also known that multi-exposure techniques have historically focused on outdoor scenes or a combination of both indoor and outdoor in the same scene and other aspects of imaging such as edges and noise, rather than stricter color accuracy which is traditionally found in art reproduction imaging [11 - 12]. However, within a complex scene such as a painting, the lower color accuracy may be within color difference tolerances and the inclusion of MEF may allow better luminance representation of the scene, particularly for darker areas as shown in earlier research works [6].

The sample of the telescope flock had consistently higher color differences than the others black samples and this may be due to it reflecting the longer wavelengths near the IR, which was found out later on after data acquisition and analysis.

Conclusion

This research assessed the color transformation accuracy of different cameras based on various targets, comparing the traditional method with target patches of extended dynamic range and with MEF imaging. The prosumer cameras generally were more accurate than the mobile phones and the addition of a wider range of luminance values in achromatic targets enhanced the color accuracy. For creating a color calibration matrix, a traditional single image (one with highest exposure without overexposure) based characterization performed better than the one created from an MEF image.

This research is part of a larger ongoing project which seeks to expand on the findings here. Currently, the project includes flatfielding, imaging of matte and glossy art, and exploring various MEF methods, and psychophysical experiments of comparison of the MEF versus the single exposure images of the art as compared to the original.

References

- [1] E. Reinhard et al, High Dynamic Range Imaging, 2nd Edition. 2010.
- [2] Shankland, Stephen. "Nokia's Lumia 1520: The First Phone That'll Take Raw Photos." CNET, CNET, 22 Oct. 2013, <https://www.cnet.com/tech/mobile/nokias-lumia-1520-the-first-phone-thatll-take-raw-photos/>.
- [3] J. R. Arillo, "Mejora del rango dinámico en la digitalización de documentos desde una perspectiva patrimonial: evaluación de métodos de alto rango dinámico (HDR) basados en exposiciones múltiples/Dynamic range improvement in document digitization from a cultural heritage perspective: assessment of high dynamic range (HDR) methods based on multiple exposures," *Revista Española De Documentación científica*, vol. 34, (4), pp. 357, 2011.
- [4] D. Marnerides, V. Hulusic and K. Debattista, "High dynamic range in cultural heritage applications," in *Visual Computing for Cultural Heritage* Anonymous Cham: Springer International Publishing, 2020, pp. 43-62.

- [5] Á. Gómez Manzanares et al, "Measuring High Dynamic Range Spectral Reflectance of Artworks through an Image Capture Matrix Hyperspectral Camera," *Sensors (Basel, Switzerland)*, vol. 22, (13), pp. 4664, 2022.
- [6] Gabrielle Brogle, "Incorporating High Dynamic Range into Multispectral Imaging for Cultural Heritage Documentation" in *Archiving Conference*, 2023, pp 94 - 96, <https://doi.org/10.2352/issn.2168-3204.2023.20.1.19>
- [7] D. Andrew Rowlands "Color conversion matrices in digital cameras: a tutorial," *Optical Engineering*, vol. 59, (11), 2020.
- [8] T. Chou and S. Chang, "Color calibration of recovering high dynamic range images," in 2008, DOI: 10.1109/CSSE.2008.1350.
- [9] F. Banterle, *Advanced High Dynamic Range Imaging: Theory and Practice*. (1st ed.) 2011. DOI: 10.1201/b11373.
- [10] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs" in 1997, DOI: 10.1145/258734.258884.
- [11] F. Xu et al, "Multi-Exposure Image Fusion Techniques: A Comprehensive Review," *Remote Sensing (Basel, Switzerland)*, vol. 14, (3), pp. 771, 2022.
- [12] T. Reiger et al, "*Technical Guidelines for Digitizing Cultural Heritage Materials: Third Edition*", Technical Guidelines, Federal Agencies Digital Guidelines Initiative, 2023.

Author Biography

Leah Humenuck is a PhD candidate in Color Science at the Munsell Color Science Laboratory at Rochester Institute of Technology. Leah's research interests are in imaging, reproduction, and lighting for cultural heritage. She is also a book and paper conservator which informs her color science research of archival items. Leah holds a BS in Chemistry from Sweet Briar College and an MA with honors in Conservation from West Dean College of Arts and Conservation.

Dr. Mekides Assefa Abebe, a computer science graduate, pursued her master's through the European Erasmus Mundus CIMET program and obtained her Ph.D. in signal and image processing via the French CIFRE program. With a rich background as a researcher in the Color and Visual Computing Laboratory at NTNU Norway, she now serves as a Visiting Assistant Professor at RIT, contributing to the Munsell Color Science Laboratory. Driven by a passion for High Dynamic Range perception and imaging, her research focuses on enhancing computer vision and graphics applications.