Beyond raw-RGB and sRGB: Advocating Access to a Colorimetric Image State

Hakki Can Karaimer[1,3] *Michael S. Brown*[1,2]
[1] *York University, Toronto* [2] *Samsung AI Centre, Toronto*
[3] *School of Computer and Communication Sciences (IC), Ecole Polytechnique Fédérale de Lausanne (EPFL)*
{karaimer, mbrown}@eecs.yorku.ca

Abstract

Most modern cameras allow captured images to be saved in two color spaces: (1) raw-RGB and (2) standard RGB (sRGB). The raw-RGB image represents a scene-referred sensor image whose RGB values are specific to the color sensitivities of the sensor's color filter array. The sRGB image represents a display-referred image that has been rendered through the camera's image signal processor (ISP). The rendering process involves several camera-specific photo-finishing manipulations intended to make the sRGB image visually pleasing. For applications that want to use a camera for purposes beyond photography, both the raw-RGB and sRGB color spaces are undesirable. For example, because the raw-RGB color space is dependent on the camera's sensor, it is challenging to develop applications that work across multiple cameras. Similarly, the camera-specific photo-finishing operations used to render sRGB images also hinder applications intended to run on different cameras. Interestingly, the ISP camera pipeline includes a colorimetric conversion stage where the raw-RGB images are converted to a device-independent color space. However, this image state is not accessible. In this paper, we advocate for the ability to access the colorimetric image state and recommend that cameras output a third image format that is based on this device-independent colorimetric space. To this end, we perform experiments to demonstrate that image pixel values in a colorimetric space are more similar across different makes and models than sRGB and raw-RGB.

Introduction

Cameras have an onboard image signal processing (ISP) module that applies a series of image manipulation steps to convert the raw-RGB sensor image to its output image. The steps applied in the ISP make up what is often referred to as the *camera imaging pipeline*. Figure 1 shows an illustrative example of the standard steps performed by an ISP. The imaging pipeline starts with a minimally processed raw-RGB image where the RGB color channels are specific to the camera sensor's spectral sensitivity. The raw-RGB image is processed through a series of steps, including steps to perform colorimetric conversion to map the raw-RGB color space to a device-independent color space–namely, one based on CIE XYZ (e.g., ProPhoto [21]). After the colorimetric conversion, additional photo-finishing steps are applied to produce a visually pleasing image which is finally encoded in a standard RGB (sRGB) color space. We can consider each of the ISP's steps as changing the state of the image, starting from a initial scene-referred color space (raw-RGB) to its final display or output-referred color space (sRGB).

Currently, users and application developers have access to only two image states: the raw-RGB image (e.g., DNG format) or the sRGB image (e.g., JPEG format). These two image states are not suitable for applications that require consistent color values across multiple cameras. For example, the raw-RGB values are specific to the sensor used by the camera. If multiple cameras observe the same scene under the same illumination, their raw-RGB values will be different due to their sensor's different spectral responses. In principle, sRGB image values should be standard across multiple devices and many applications erroneously make this assumption. However, this assumption overlooks the fact that virtually all camera pipelines include camera-specific photo-finishing operations that modify the image values to produce a visually pleasing image. Moreover, the type of photo-finishing applied is often specific to the camera's settings during imaging, such as the picture style (e.g., landscape style, vivid style, portrait style). As a result, sRGB values not only are dissimilar among different cameras but also can even differ on the same camera when different capture settings are used.

Contribution In this paper, we advocate for a third image format to be made available to consumers and application developers. Specifically, we argue that the internal ISP image state after it has been processed by the colorimetric steps in the camera pipeline is more suitable for applications that assume imaged scene values are consistent among multiple cameras. To demonstrate this, we performed experiments that analyze the consistency of pixel values at different image states for scene points imaged by multiple DSLR cameras. We show that the colorimetric image state is by far the most consistent across multiple cameras. These experiments reinforce our advocacy for camera manufacturers to provide access to a colorimetric image format.

Camera ISP Preliminaries

We begin with a quick overview of the camera processing pipeline shown in Figure 1. This figure illustrates several of the high-level steps performed by a typical ISP. Readers are referred to [16, 20] for more details about standard camera pipelines.

The camera pipeline starts with the raw-RGB image that represents a minimally processed image captured from the camera's sensor. The raw-RGB values can be considered scene-referred as they are directly related to the image scene; however, the values are in a color space specific to the spectral sensitivities of the sensor's color filter array. The pipeline can be divided into three stages. The first stage involves raw-RGB processing that includes routines such as linearizing the raw-RGB values and correcting for issues related to the lens' vignetting and chromatic abberations. The raw-RGB image is also demosiaced to provide three full-color channels. The pipeline second stage involves a colorimetric conversation that applies routines to convert the raw-RGB values to device-independent values based on the CIE 1931 color space. Note that this includes a white-balance step realized as a 3 $\times$ 3 diagonal matrix operation followed by a linear colorimetric step as a 3 $\times$ 3 matrix operation. Most cam-

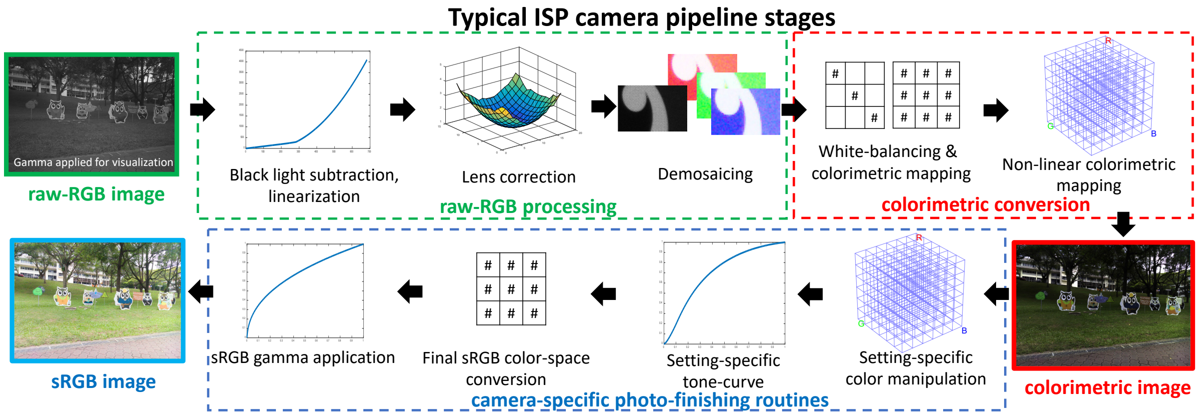

Figure 1. *This figure (adapted from [16]) shows the standard steps applied on a camera's ISP. Currently users have access to the raw-RGB image and the sRGB image. The pipeline can be divided into three stages: (1) raw-RGB image processing; (2) a colorimetric conversion to map the raw-RGB values to device-independent values based on CIE XYZ; (3) camera-specific photo-finishing. We advocate access to the colorimetric image state.*

eras include an additional nonlinear colorimetric step that is performed as a 3D lookup table (LUT). After this stage, the image values are in a colorimetric color space that is directly related to the scene (i.e., still a scene-referred color space). After the colorimetric conversion, the final stage applies photo-finishing operations to modify the image further to be visually appealing and to be suitable for use on display devices. The final image is encoded in a display-referred sRGB color space.

Related Work

There are two areas related to the work addressed in this paper. First are methods targeting camera colorimetric calibration. There are number of papers that focus on mapping the raw-RGB values to a perceptual color space (e.g., [1, 2, 8–10, 14]). These methods typically focus on how best to compute an accurate mapping between raw-RGB and CIE 1932 XYZ values by imaging a physical calibration chart with known CIE XYZ values. Most work focuses on the types of mathematical functions used to parameterize the color mappings, such as high-order polynomials (e.g., [10, 14]), look-up-tables (e.g., [15]), or neural networks (e.g., [4]). Issues such as non-uniform illumination are also addressed (e.g., [1, 8, 9]) as it can be difficult to ensure uniform illumination on the physical calibration pattern. There has even been work that demonstrates that additional physical filters achieve better colorimetric properties (e.g., [7, 11]). Camera manufacturers have undoubtedly incorporated these research findings into their camera pipelines; however, users do not have access to such results unless they manually perform and apply the calibration methods discussed in these works.

Radiometric calibration (sometimes called color de-rendering) is another related research area. Radiometric calibration is the process of modeling the camera's imaging pipeline in order to undo the photo-finishing routines applied. This is often performed when scene-referred values are needed, but only sRGB images are available. Notable works include [3, 5, 6, 12, 13, 17–19, 22]. These works propose mathematical models to emulate the steps of the camera pipeline. Once a model is established, it can be inverted to obtained image values at various states in the image pipeline. The work in this paper advocates a third output format that encodes the colorimetric values in the *existing* ISP camera pipeline, thus avoiding the need for radiometric calibration.

Experimental Setup and Analysis

As described in the previous section, the goal of this paper is to demonstrate that providing access to image values after the colorimetric image conversion is the most suitable image state to provide consistent image values among different cameras.

Towards this goal, we perform the following experiments illustrated in Figure 2. The following four DSLR cameras are used to capture scenes under various lighting conditions: (1) a Canon 5D Mark IV, (2) a Nikon D7200, (3) a Fujifilm X-T20, and (4) an Olympus OM-D E-M10. These cameras are placed in a controlled environment where each camera captures three types of planar scenes with different physical materials. Specifically, we image a Macbeth Color Rendition Chart (with 24 color patches), and two charts of 81 paint samples that we have created, denoted as Paint Samples #1 and Paint Samples #2. The scenes are illuminated by a tunable direct current (DC) lighting rig. We use illuminations that correspond to correlated color temperatures of 3200°K, 4340°K, and 5500°K.

Each scene and lighting condition are captured with the four DSLR cameras and saved in a raw-RGB image format. For these particular cameras, the raw-RGB image format embeds colorimetric conversion metadata that provides the parameters for the linear transformation and nonlinear 3D LUTs. Using the software camera pipeline from Karaimer and Brown [16] we are able to render each state of the imaging pipeline. To avoid errors due to each camera's proprietary auto-white-balance functions, we perform the white-balance operation based on the neutral color patches in the Macbeth color rendition chart. Our imaging setup is shown in Figure 2-(A).

In order to assess the consistency among the cameras, we map each observed patch's pixel values to 2D chromaticity space as shown in Figure 2-(B). We use the 2D chromaticity space to factor out differences related to overall image brightness among the cameras due to issues such as image gain and differences in exposure. We compute the mean variance of each patch in the 2D chromaticity space for the following image states: (1) raw-RGB, (2) raw-RGB with white balance applied, (3) raw-RGB with the linear colorimetric operation applied (4) raw-RGB with the additional non-linear (3D LUT) colorimetric operation applied, and (5) sRGB.

The results for the Macbeth Color Rendition Chart, Paint Samples #1, and Paint Samples # 2 are shown in Table 1. The table summarizes the results for each CCT as well as the overall

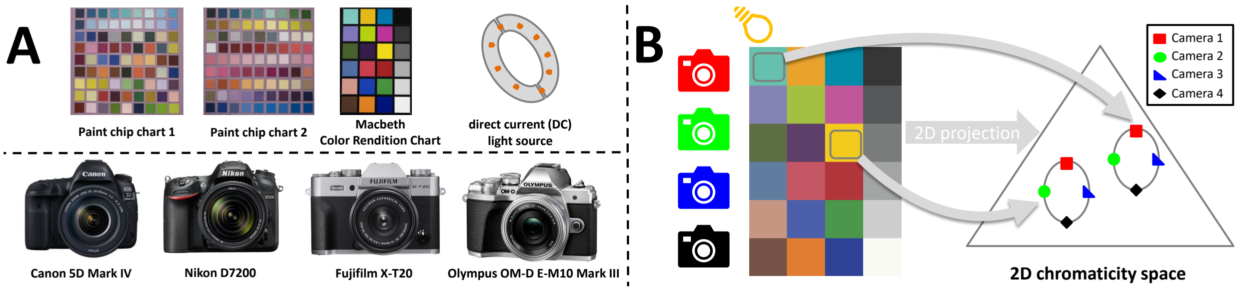

Figure 2. (A) Our experiment setup. Patches from three different physical patterns (Macbeth Color Rendition Chart, and two custom charts of 81 flat paint samples each) are imaged using four different cameras. A direct current (DC) light source is used to illuminate the scene under different correlated color temperatures (CCTs). (B) For each color patch observed by the four cameras, we compare their values in a 2D chromaticity space. In particular, we are interested in the variance of these measures in different internal camera pipeline image states.

(1) Color Rendition Chart				
Stage	3200°K	4340°K	5500°K	Combined
raw-RGB	0.046	0.049	0.051	0.130
white-balanced raw-RGB	0.018	0.017	0.018	0.016
colorimetric image state (linear)	0.125	0.117	0.103	0.095
colorimetric image state (non-linear)	**0.004**	**0.004**	**0.003**	**0.004**
sRGB	0.041	0.041	0.042	0.037

(2) Paint Samples #1				
Stage	3200°K	4340°K	5500°K	Combined
raw-RGB	0.121	0.131	0.139	0.411
white-balanced raw-RGB	0.031	0.028	0.031	0.029
colorimetric image state (linear)	0.213	0.174	0.147	0.148
colorimetric image state (non-linear)	**0.006**	**0.004**	**0.005**	**0.009**
sRGB	0.042	0.029	0.038	0.041

(3) Paint Samples #2				
Stage	3200°K	4340°K	5500°K	Combined
raw-RGB	0.127	0.135	0.143	0.424
white-balanced raw-RGB	0.034	0.033	0.035	0.031
colorimetric image state (linear)	0.218	0.177	0.155	0.152
colorimetric image state (non-linear)	**0.004**	**0.004**	**0.005**	**0.007**
sRGB	0.045	0.049	0.056	0.051

Table 1: Colorimetric consistency between four cameras at five different processing stages of the camera imaging pipeline using the patches from (1) the Macbeth Color Rendition Chart, (2) our first custom chart with 81 paint samples, (3) our second custom chart with 81 different paint samples. The table reports the average variance of each patch's 2D chromaticity values among the four cameras in the respective image states. The most consistent method is highlighted in bold and green. Results are shown for each illumination CCT and all illuminations combined.

result where all illumination conditions are combined. It is clear that the full colorimetric conversion image state (raw-RGB with linear and non-linear colorimetric conversion applied) is by far the most consistent state.

Figures 3–5 show the plots in chromaticity space for the three imaged charts. Each cluster of points represents the same patch imaged by the four cameras. An ellipse is shown to show the spread of the measurements. From these plots, we can also see that the full colorimetric conversion image state has the minimum spread (i.e., lowest variance among the cameras).

Discussion Our experiments show that the colorimetric mapping stage achieves the most consistent results. More importantly, we obtain these results without the need for any colorimetric calibration; this is simply a matter of accessing the appropriate image state within the existing camera imaging pipeline. As a result, we advocate the need for cameras to allow a third image format beyond raw-RGB and sRGB, named a colorimetric image state.

Concluding Remarks

We have performed experiments to demonstrate the advantages of allowing access to the colorimetric image state in the camera imaging pipeline. Our experiments showed that image values in the colorimetric image state have the least amount of variance among different cameras when compared to other image states. Currently only the raw-RGB and sRGB image states are available to users, which represent arguably the two *worst* image states for multi-camera consistency. We point out that the white-balanced raw-RGB image state included in our experiments was provided only for comparison; like the colorimetric image state, the white-balance raw-RGB image state is not easily accessible.

It is worth noting that our experiments compared the variance of values in different images states that represent different color spaces. This is important to consider as variance is not invariant to scaling which undoubtedly happens in the conversion between color spaces. A more proper evaluation would involve attempting to normalize each color space into a canonical 3D volume and then estimate the variance. However, such normalization does not represent how image processing and computer vision application developers use such values, nor how the camera pipeline manipulates the image's RGB values. Instead, applications developers would tune their algorithms based on the color space used, and in such cases, having lower variance among different devices is always preferred.

Finally, our experiments were limited to DSLR cameras, as smartphone cameras currently do not include the nonlinear colorimetric metadata in their raw-RGB files. This lack of metadata made it impossible to evaluate the true colorimetric conversion performed onboard smartphone cameras. This issue further bolsters our advocacy for more open access to various image states within the camera image pipeline. In this paper, we show compelling evidence that accessing the colorimetric image is useful; however, access to all image states would be welcomed. Analysis of other image states for both DSLR and smartphone cameras is part of our ongoing and future efforts.

Acknowledgments This study was funded in part by a Google Faculty Research Award, the Canada First Research Excellence Fund for the Vision: Science to Applications (VISTA) programme, an NSERC Discovery Grant. Dr. Brown contributed to this article in his personal capacity as a professor at York University. The views expressed are his own and do not necessarily represent the views of Samsung Research.

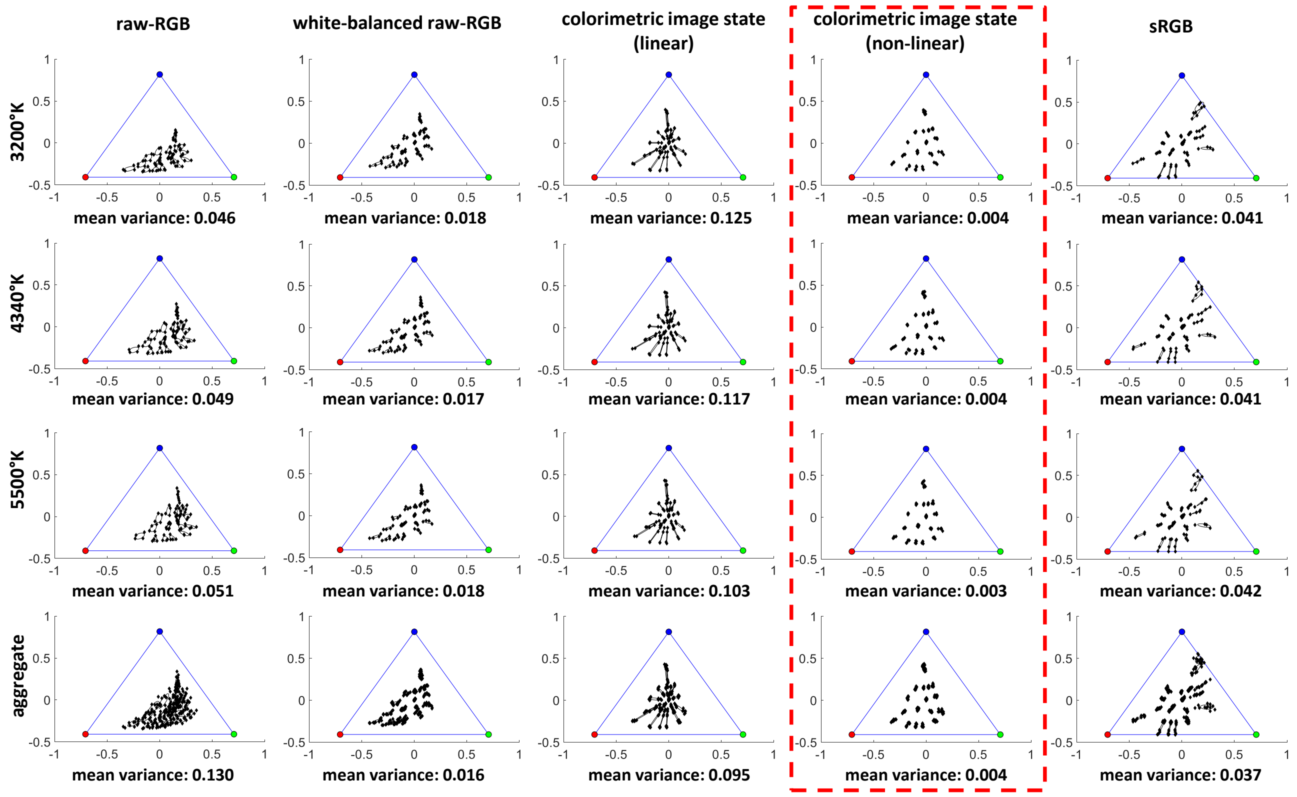

Figure 3. *Colorimetric consistency between four cameras at five different processing stages of the camera imaging pipeline using the patches of the Macbeth Color Rendition Chart. It is clear that the full colorimetric conversion image state (raw-RGB with linear and non-linear colorimetric conversion applied) is by far the most consistent state.*

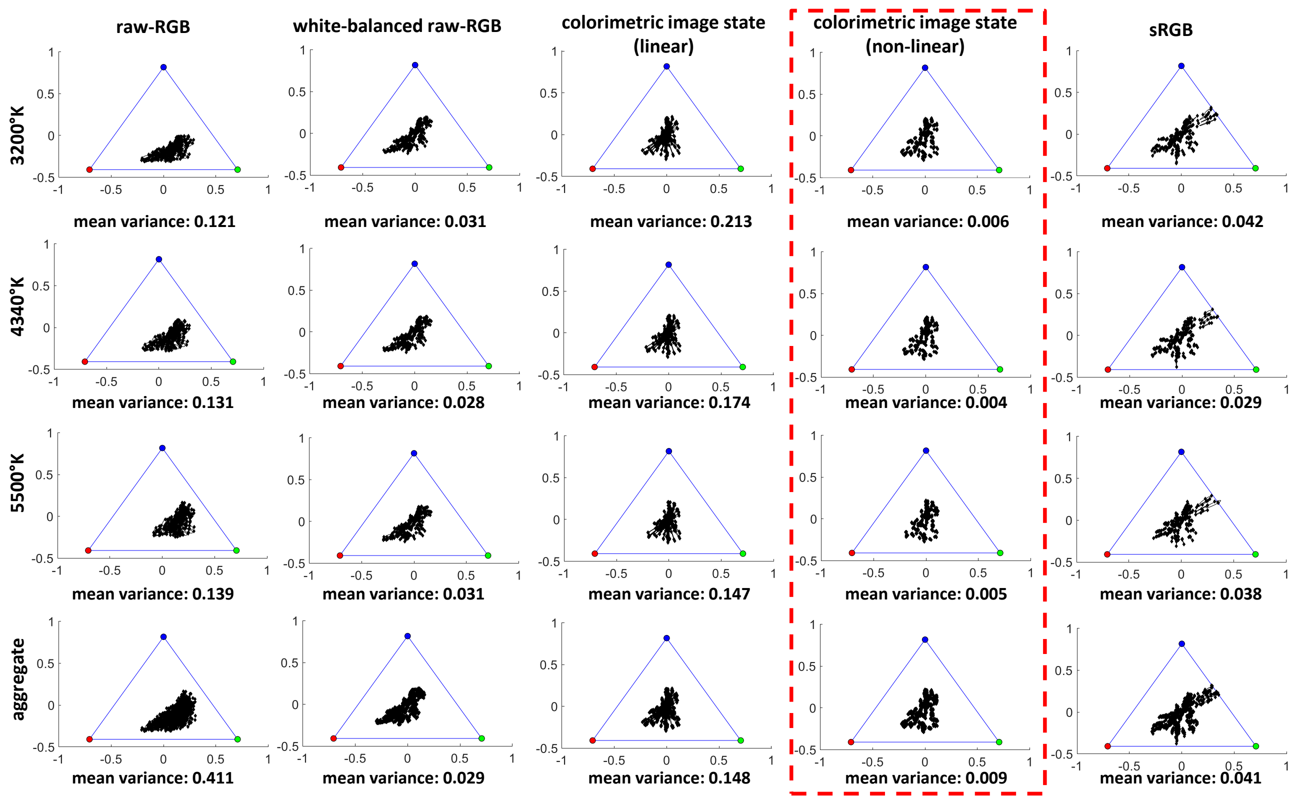

Figure 4. *Colorimetric consistency between four cameras at five different processing stages of the camera imaging pipeline using the patches of the 81 Paint Samples #1. It is clear that the full colorimetric conversion image state (raw-RGB with linear and non-linear colorimetric conversion applied) is by far the most consistent state.*

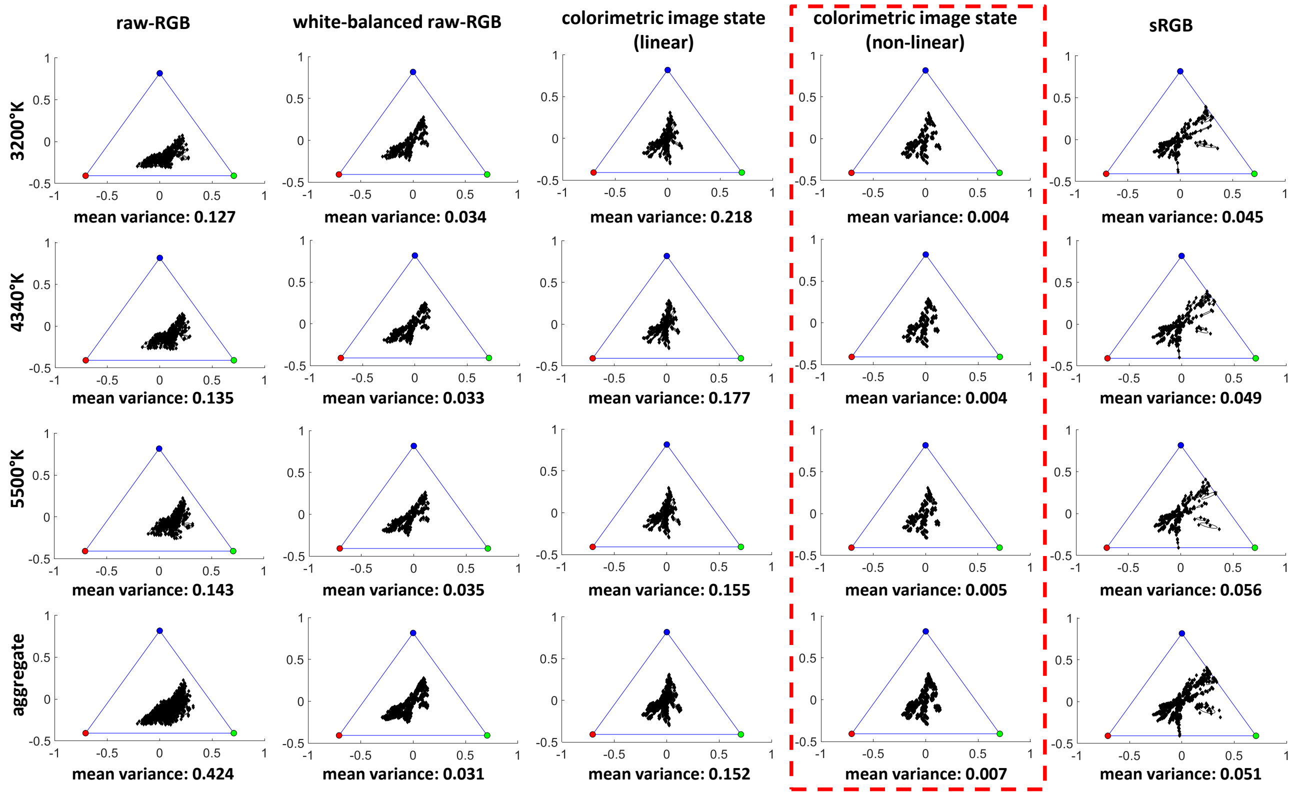

Figure 5. *Colorimetric consistency between four cameras at five different processing stages of the camera imaging pipeline using the patches of the 81 Paint Samples #2. It is clear that the full colorimetric conversion image state (raw-RGB with linear and non-linear colorimetric conversion applied) is by far the most consistent state.*

References

[1] P. Bastani and B. Funt. Simplifying irradiance independent color calibration. In *Color and Imaging Conference*, 2014.

[2] S. Bianco, A. Bruna, F. Naccari, and R. Schettini. Color-space transformations for digital photography exploiting information about the illuminant estimation process. *Journal of Optical Society America A*, 29(3):374–384, 2012.

[3] A. Chakrabarti, D. Scharstein, and T. Zickler. An empirical camera model for internet color vision. In *BMVC*, 2009.

[4] T. L. V. Cheung and S. Westland. Color camera characterisation using artificial neural networks. In *Color and Imaging Conference*, 2002.

[5] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, 1997.

[6] M. Diaz and P. Sturm. Radiometric calibration using photo collections. In *ICCP*, 2011.

[7] J. E. Farrell and B. A. Wandell. Method and apparatus for identifying the color of an image. U.S. Patent No. 5,479,524, Washington, DC: U.S. Patent and Trademark Office, 1995.

[8] G. D. Finlayson, H. Gong, and R. B. Fisher. Color homography color correction. In *Color and Imaging Conference*, 2016.

[9] G. D. Finlayson, M. M. Darrodi, and M. Mackiewicz. The alternating least squares technique for nonuniform intensity color correction. *Color Research & Application*, 40(3):232–242, 2015.

[10] G. D. Finlayson, M. Mackiewicz, and A. Hurlbert. Color correction using root-polynomial regression. *IEEE Transactions on Image Processing*, 24(5):1460–1470, 2015.

[11] G. D. Finlayson, Y. Zhu, and H. Gong. Using a simple colour prefilter to make cameras more colorimetric. In *Color and Imaging Conference*, 2018.

[12] H. Gong, G. D. Finlayson, and M. M. Darrodi. Concise radiometric calibration using the power of ranking. In *BMVC*, 2017.

[13] H. Gong, G. D. Finlayson, M. M. Darrodi, and R. B. Fisher. Rank-based radiometric calibration. In *Color and Imaging Conference*, 2018.

[14] G. Hong, M. R. Luo, and P. A. Rhodes. A study of digital camera colorimetric characterisation based on polynomial modelling. *Color Research & Application*, 26(1):76–84, 2001.

[15] P.-C. Hung. Colorimetric calibration in electronic imaging devices using a look-up-table model and interpolations. *Journal of Electronic Imaging*, 2(1):53–61, 1993.

[16] H. C. Karaimer and M. S. Brown. A software platform for manipulating the camera imaging pipeline. In *ECCV*, 2016.

[17] S. J. Kim, H. T. Lin, Z. Lu, S. Süsstrunk, S. Lin, and M. S. Brown. A new in-camera imaging model for color computer vision and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(12):2289–2302, 2012.

[18] S. Nam and S. J. Kim. Modelling the scene dependent imaging in cameras with a deep neural network. In *ICCV*, 2017.

[19] R. M. Nguyen and M. S. Brown. Raw image reconstruction using a self-contained srgb-jpeg image with only 64 kb overhead. In *CVPR*, 2016.

[20] R. Ramanath, W. E. Snyder, Y. Yoo, and M. S. Drew. Color image processing pipeline. *IEEE Signal Processing Magazine*, 22(1):34–43, 2005.

[21] K. E. Spaulding, E. Giorgianni, and G. Woolfe. Reference input/output medium metric rgb color encodings (rimm/romm rgb). In *Image Processing, Image Quality, Image Capture, Systems Conference*, 2000.

[22] Y. Xiong, K. Saenko, T. Darrell, and T. Zickler. From pixels to physics: Probabilistic color de-rendering. In *CVPR*, 2012.