Assessing Color Reproducibility of Whole-Slide Imaging Scanners

Wei-Chung Cheng, Tyler Keay, Neil O'Flaherty, Joel Wang,
Adam Ivansky, Marios A. Gavrielides, Brandon D. Gallas, Aldo Badano
Division of Imaging and Applied Mathematics, CDRH, U.S. Food and Drug Administration
10903 New Hampshire Avenue, Silver Spring, MD USA 20993-0002

ABSTRACT

A new method for assessing color reproducibility of whole-slide imaging (WSI) systems is introduced. A color phantom is used to evaluate the difference between the input to and the output from a WSI system. The method consists of four components: (a) producing the color phantom, (b) establishing the truth of the color phantom, (c) retrieving the digital display data from the WSI system, and (d) calculating the color difference. The method was applied to a WSI system and used to evaluate the color characteristics with and without color management.

Keywords: Digital pathology, whole-slide imaging, color reproducibility, color phantom, color management, ICC profile

1. INTRODUCTION

The goal of this study is to evaluate the color reproducibility of a WSI scanner, which is determined by the color difference between the ground truth of the input slide and the output image. Neither determining the ground truth of the input slide nor measuring the output image on the display is straightforward.

Establishing the color truth of the input slide can be challenging. The color stimulus of a tissue slide, as a transmissive object, is generated by the interaction between the lighting and the transmittance of the object. Although the tissue transmittance is invariable, the light source needs to be fixed to determine the color. Unfortunately, the light source in optical microscopes such as a halogen-tungsten lamp is not standardized and should not be used as part of the truth. In addition, the cellular structures in tissue slides are too small for any colorimeter to select a uniform spot to measure. Thus color phantoms are required to test the WSI systems.

On the display side, the scanned image of the WSI system is reproduced by the display sub-system, which consists of the factory review software, computer software/hardware, and display device. The factory review software is responsible for parsing the WSI image file, imposing optional image enhancement, and activating the color management options. Therefore the review software should be included in the color assessment for its important role. However, the display device is a swappable component and usually introduces considerable variability due to calibration errors and user settings. Although the complexity of the display device calls for separate characterization and modeling, fortunately the display interface always uses the standard sRGB color space and provides a robust digital data output of the WSI system.

In 2011, Dr. Yukako Yagi at Massachusetts General Hospital proposed a method for assessing WSI systems [1]. In the study, two types of color phantoms were proposed—slide color phantoms and digital color phantoms. The slide color phantoms were designed for hematoxylin and eosin (H&E) staining and comprised 9 color patches made of color filters. The size and arrangement were designed to fit a 4x objective. The digital color phantoms, which were image files obtained by scanning the slide color phantoms, were used to evaluate the variation among 23 displays of the same model in the Pathology Department. A display analyzer was used to read the RGB/HSL values of the color patches from the display. The results showed very pronounced inter-display variation. Although this study effectively advocated the importance of color reproducibility of WSI systems, the proposed method focused only on the consistency of the image display sub-system and did not address the image acquisition subsystem quantitatively.

In this study, we developed a new method of establishing the color truth with a color phantom for assessing color reproducibility of WSI systems. In our approach, the color phantom is made of photographic film with 140 color patches mounted on a blank glass slide. The truth is based on the measurement and calculation of the spectral transmittance of each individual color patch on the color phantom. For the WSI output, the digital display data retrieved from the display interface are used to represent the output of the WSI system. Finally, standard CIE formulas are used to calculate the colorimetrical differences.

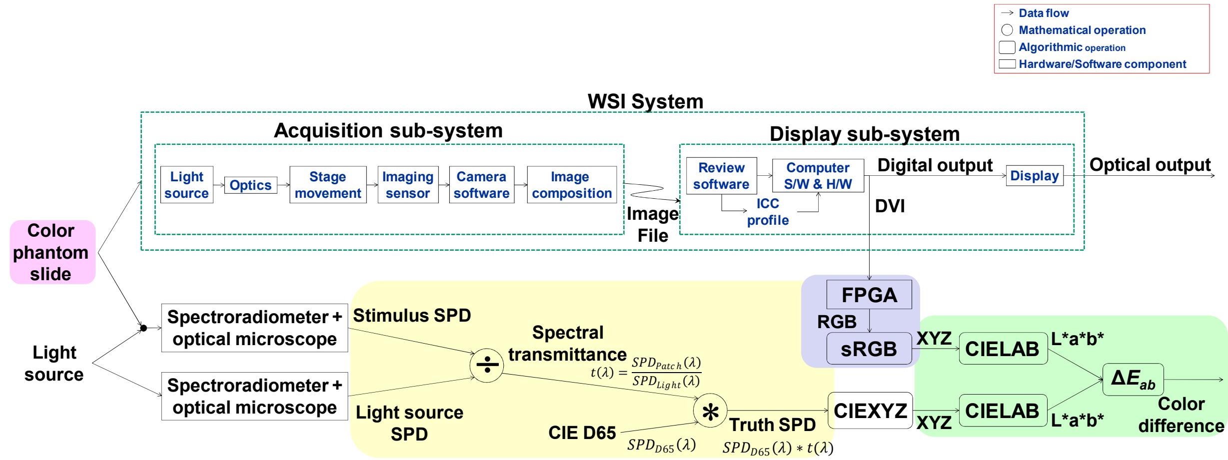

Figure 1. Data flow of the methodology. The color phantom slide goes through the upper stream, the WSI system, to obtain the digital display data, which are intercepted by an FPGA circuit board as the outcome. In the lower stream, the phantom is illuminated by a stable light source for measuring and calculating the spectral transmittance. The truth is defined as the product of the spectral transmittance and the standard CIE D65 illuminant. Finally the CIELAB color difference is calculated to indicate the color reproducibility of the WSI system.

The advantages of our approach can be summarized as follows. Totally 140 or more color patches can be chosen to represent the color gamut of interest. The color truth is independent of the lighting so any stable light source can be used in this method. Reading the digital display data is robust and eliminates the time-consuming optical measurement. The final evaluation is reported in the universal, device-independent CIE color difference and suitable for inter-WSI system comparison.

2. METHOD

2.1 Producing color phantoms

For evaluating a WSI system, the color phantom must satisfy three criteria: (a) it must be transparent and of acceptable size for the WSI scanner, (b) the selection of the color patches must be broad enough to be representative of various stains and tissue types, (c) the size of each color patch must be large enough for the instruments to measure, while small enough such that all color patches can be scanned by the WSI scanner at once to reduce intra-scanner variability. We used photographic transparency film (Fujichrome Velvia 100, Fuji Photo Film Co., Ltd., Tokyo, Japan) to photograph the color target that consists of 14x10 color patches (GretagMacbeth ColorChecker SG, X-Rite Inc., MI, USA). The specific film was chosen because it offers the finest grain in the market. Fine grain helps reduce the non-uniformity observed under high power objectives. The color target was illuminated by daylight-matching fluorescent lamps in a light booth (ColorMatcher 6500K lamps and PDV-2e, GTI Graphic Technology, Inc., Newburgh, NY, USA). The photographs were taken with a legacy film camera (Nikon F3 HP and Nikkor 50mm f/1.4 AIS, Nikon Corp., Tokyo, Japan). A wide range of exposure settings were swept to compensate for the potential mechanical and chemical variation in the uncalibrated camera shutter, lens aperture, film, and lab processing. After processing by a professional lab (Dwayne's Photo, Parsons, Kansas) with the standard E-6 process, a well-exposed 24x36 mm^2 frame, which properly rendered the 140 color patches, was chosen, trimmed and attached to a blank 1"x3" glass slide with its emulsion side facing up as shown in Figure 2. Every color patch is large enough to cover the field of view of a 20x microscope, i.e., 1.15 mm in diameter. Notice that this color phantom does not necessarily reproduce the original scene (i.e., the ColorChecker).

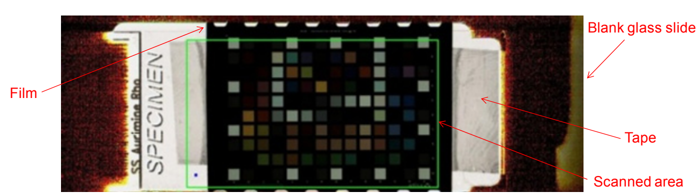

Figure 2. A snapshot of the color phantom taken by the WSI scanner.

2.2 Establishing the color truth

Establishing the color truth of the transparent color patches consists of two steps – determining the spectral transmittance of each color patch followed by calculating the truth. A spectroradiometer with a flexible probe (PR730 and FP730, Photo Research, Inc., Chatsworth, CA, USA) was installed in a microscope with a 20x objective and a tungsten-halogen lamp (AxioPlan2, A-Plan 20x, and HAL 100, Carl Zeiss Microscopy, LLC, Thornwood, NY, USA). The spectral power distribution (SPD) of each color patch was measured automatically with a motorized XY-stage (MAC5000, Ludl Electronic Products Ltd., Hawthorne, NY, USA).

The spectral transmittance of each color patch, $t(\lambda)$, was calculated by dividing the measured stimulus SPD of each color patch, $SPD_{Patch}(\lambda)$, by the SPD of the light source, $SPD_{Light}(\lambda)$

$$t(\lambda) = \frac{SPD_{Patch}(\lambda)}{SPD_{Light}(\lambda)}. \tag{1}$$

Then the SPD of standard CIE D65 illuminant, $SPD_{D65}(\lambda)$, was applied to the spectral transmittance to obtain the truth, $SPD_{Truth}(\lambda)$

$$SPD_{Truth}(\lambda) = SPD_{D65}(\lambda) * t(\lambda). \tag{2}$$

The truth SPD (or any other SPD) was converted into the CIEXYZ color space as a tristimulus, X, Y and, Z, by

$$X = \int_{380}^{780} SPD(\lambda) * \bar{x}(\lambda)\, d\lambda \tag{3a}$$

$$Y = \int_{380}^{780} SPD(\lambda) * \bar{y}(\lambda)\, d\lambda \tag{3b}$$

$$Z = \int_{380}^{780} SPD(\lambda) * \bar{z}(\lambda)\, d\lambda \tag{3c}$$

where $\bar{x}$, $\bar{y}$ and $\bar{z}$ are the CIE 1964 color matching functions defined as numerical data [2] and depicted in Figure 3.

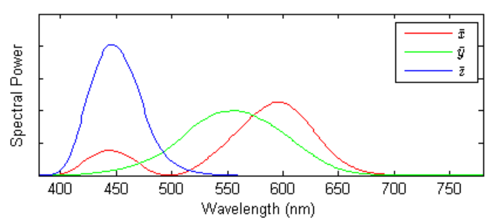

Figure 3. The CIE 1964 color matching functions.

The truth tristimulus was then converted into the CIELAB color space

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16 \tag{4a}$$

$$a^* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \tag{4b}$$

$$b^* = 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \tag{4c}$$

$$f(t) = \begin{cases} t^{\frac{1}{3}}, & t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29}, & t \le \left(\frac{6}{29}\right)^3 \end{cases} \tag{4d}$$

where X_n, Y_n, and Z_n represent the tristimulus values of the reference white. For the color phantom slide, the reference white was defined by the white colors from the 140 color patches. The CIELAB color space accounts for the light adaptation phenomenon via the normalization to the reference white. The CIELAB conversion is also a necessary step for calculating the color difference described in Section 2.4.

2.3 Retrieving the WSI output

The color phantom was scanned by a WSI system (ScanScope CS, Aperio, Vista, CA, USA). Since the color phantom was substantially different from regular tissue slides in terms of texture, size, and flatness, unsurprisingly the WSI scanner failed to focus automatically. Thus the manual focus mode was used to select more than 20 focus points across the whole area. A blank area outside the color phantom (i.e., solely the blank glass slide) was used for the scanner to execute the flat-field correction. A white color patch (i.e., a transparent piece of the film on the blank glass slide) was used to execute the white balancing. The output file was encoded in the proprietary SVS format with an embedded ICC color profile. The SVS files can only be decoded and reviewed by the factory review software (Aperio ImageScope), in which the ICC color management could be optionally activated or deactivated. The final image was reproduced on the monitor via the Digital Video Interface (DVI) interface.

A custom-designed circuit was used to record the pixel data of the final image. The circuit was implemented in a field programmable gate array (FPGA) board, which was inserted between the computer and display to intercept the RGB values of a predefined pixel in the center of the display [3].

For each color patch, 10 samples in different locations were taken and averaged to determine the color. Like most display systems, the WSI system assumed the display using the standard sRGB color space [4]. The sRGB is a well-defined, device-independent color space and can be converted into the CIEXYZ color space by

$$C_{linear} = \begin{cases} \frac{C_{srgb}}{12.92}, & C_{srgb} \leq 0.03928 \\ \left(\frac{C_{srgb}+0.055}{1.055}\right)^{2.4}, & C_{srgb} > 0.03928 \end{cases} \tag{5a}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} \tag{5b}$$

where C represents either R, G, or B. Then the WSI output color was converted from CIEXYZ to CIELAB.

2.4 Calculating the color differences

The CIELAB ΔE^*_{ab} color difference between the WSI output color (L^*_1, a^*_1, b^*_1) and the color truth (L^*_2, a^*_2, b^*_2) was calculated by

$$\Delta E^*_{ab} = \sqrt{(L^*_2 - L^*_1)^2 + (a^*_2 - a^*_1)^2 + (b^*_2 - b^*_1)^2}. \tag{6}$$

3. RESULTS

The CIELAB ΔE^*_{ab} color differences of the 140 color patches are shown in Fig. 4. Without color management, the ΔE^*_{ab} ranges between 0.53 and 36.46 with μ=13.60 and σ=9.74. With color management, the range was reduced to 0.51 and 30.26 with μ=9.58 and σ =7.98. The color differences of 118 color patches were reduced by color management. The results are considered substandard compared with imaging systems that are designed for reproducing color faithfully such as photo scanners.

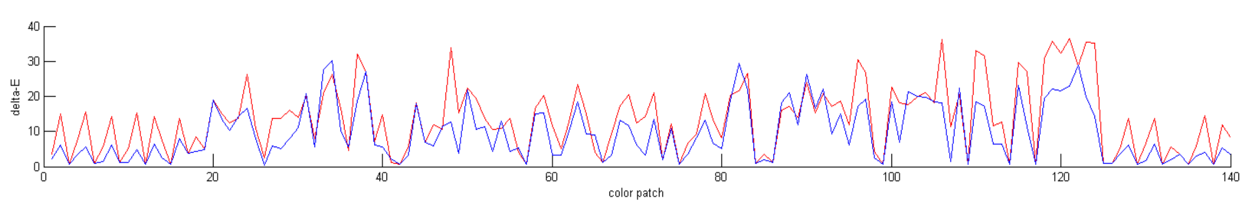

Figure 4. CIELAB 1976 color differences of the 140 color patches between the truth and the scan. The blue and red curves indicate the color management-activated and -deactivated results, respectively.

4. CONCLUSIONS

A standard, objective, device-independent colorimetric method was introduced to assess the color reproducibility of a WSI system. A photographic technique was used to fabricate the color phantom. The definition of the truth was based on the invariable spectral transmittance of the color phantom. The technique of retrieving the digital display data with a circuit board was described. The color differences were calculated with the standard CIE formulas.

The method was applied to a WSI system and revealed its color characteristics. The effects of the ICC color profile were also evaluated. The results show pronounced color differences for certain hues when color management is activated and even worse performance without color management. The findings suggest that color reproducibility of WSI scanners demands careful adjustments within a color management framework.

Although producing and measuring the color phantoms require special instrumentation and skills, these tasks can be commissioned to a professional lab that is properly equipped to deliver the color phantoms accompanied with accurate measurement data. Another device required by this method is the FPGA board, which is commercially available for under $500 USD.

For future work, the minimum color difference required for performing specific diagnostic tasks will be investigated.

ACKNOWLEDGEMENTS

The authors thank Dr. Stephen Hewitt (NIH/NCI) for loaning the optical microscope system and Drs. Rodney Rouse and Katherine Shea (FDA/CDER) for sharing the WSI system. This study was supported in part by a grant from the Critical Path Initiative at the Food and Drug Administration and by an appointment to the Research Participation Program at CDRH administered by ORISE through an interagency agreement between DOE and FDA. The mention of commercial products herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services. This is a contribution of the Food and Drug Administration and is not subject to copyright.

REFERENCES

[1] Yagi, Y., "Color standardization and optimization in whole slide imaging," Diagnostic Pathology 6(Suppl 1):S15, (2011). www.diagnosticpathology.org/content/6/S1/S15

[2] Wyszecki, G. and Stiles, W. S., [Color Science: Concepts and Methods, Quantitative Data and Formulae], John Wiley & Sons, Inc., New York, (1982).

[3] Cheng, W.-C. and Badano, A., "Virtual Display: A Platform for Evaluating Display Color Calibration Kits," SID Int. Symp. Digest Tech. Papers **42**, 1030–1033, (2012).

[4] IEC 61966-2-1, "Multimedia systems and equipment – Colour measurement and management – Part 2-1: Colour management – Default RGB colour space – sRGB" (1999).